

# The CMB Bispectrum

J.R. Fergusson, M. Liguori, and E.P.S. Shellard

*Centre for Theoretical Cosmology,  
Department of Applied Mathematics and Theoretical Physics,  
University of Cambridge, Wilberforce Road,  
Cambridge CB3 0WA, United Kingdom*

(Dated: June 9, 2010)

We use a separable mode expansion estimator with WMAP data to estimate the bispectrum for all the primary families of non-Gaussian models, including non-scaling feature models, the flattened model, DBI and ghost inflation, as well as previously constrained simple cases. We review the late-time mode expansion estimator methodology which can be applied to any non-separable primordial and CMB bispectrum model, and we demonstrate how the method can be used to reconstruct the CMB bispectrum from an observational map. We extend the previous validation of the general estimator using local map simulations. We apply the estimator to the coadded WMAP 5-year V and W channel maps, reconstructing the WMAP bispectrum using  $l < 500$  multipoles and  $n = 31$  orthonormal 3D eigenmodes; both the mode expansion parameters and the reconstructed 3D WMAP bispectrum are plotted. We constrain all popular nearly scale-invariant models, ensuring that the theoretical bispectrum is well-described by a convergent mode expansion. Constraints from the local model  $f_{\text{NL}} = 54.4 \pm 29.4$  and the equilateral model  $f_{\text{NL}} = 143.5 \pm 151.2$  ( $F_{\text{NL}} = 25.1 \pm 26.4$ ) are consistent with previously published results. (Here, we use a nonlinearity parameter  $F_{\text{NL}}$  normalised to the local case, to allow more direct comparison between different models.) Notable new constraints from our method include those for the constant model  $F_{\text{NL}} = 35.1 \pm 27.4$ , the flattened model  $F_{\text{NL}} = 35.4 \pm 29.2$ , and warm inflation  $F_{\text{NL}} = 10.3 \pm 27.2$ . We investigate feature models, which break scale invariance, surveying a wide parameter range for both the scale and phase, and we find no significant evidence of non-Gaussianity for all cases well-described by the given eigenmodes (i.e. with features  $l^* > 150$ ). We propose a measure  $\bar{F}_{\text{NL}}$  for the total integrated bispectrum and find that the measured value is consistent with the null hypothesis that CMB anisotropies obey Gaussian statistics. We argue that this general bispectrum survey with the WMAP data represents the best evidence for Gaussianity to date and we discuss future prospects with higher precision and resolution, notably from the Planck satellite.

## I. INTRODUCTION

In an earlier paper [1], we described a general approach to the estimation of non-separable CMB bispectra using separable mode expansions. Our aim here is to directly estimate the full CMB bispectrum from WMAP data, to survey and constrain current non-Gaussian primordial theories, and to discuss the prospects for reconstructing the bispectrum with forthcoming data, such as the Planck experiment. Previous work by other groups has endeavoured to measure the bispectrum by using specific estimators tailored to investigate particular separable models, such as the well-known local and equilateral bispectra. This restriction to separable cases was for reasons of calculational simplicity to make the data analysis tractable, that is, reducing it from  $\mathcal{O}(l_{\text{max}}^5)$  to  $\mathcal{O}(l_{\text{max}}^3)$  operations. We summarise constraints that have been obtained to date using these methods later in section V, when we survey theoretical models; it is sufficient at this point to note that the present WMAP7 constraint  $-10 < f_{\text{NL}} < 74$  [2] (95% confidence) does not provide any significant evidence for a primordial local bispectrum signal, and nor do constraints on the few other models investigated to date (see the review [3]).

Two significant developments mean that we can move beyond these specific estimators and consider a more general approach which includes the reconstruction of the whole bispectrum directly from the observational data. First, explicit calculations of the reduced CMB bispectrum  $b_{l_1 l_2 l_3}$  in a wide-ranging survey of primordial theories [4, 5], demonstrated that the resulting coherent patterns of acoustic peaks could be represented by rapidly convergent mode expansions with a limited number of terms (irrespective of whether the primordial bispectrum was separable). Secondly, these complete orthonormal mode expansions could be transformed into a non-orthogonal frame with separable basis functions [1] in which the same simplifications could be exploited to efficiently calculate the estimator (11) in  $\mathcal{O}(l_{\text{max}}^3)$  operations, again for arbitrary non-separable theoretical bispectra  $b_{l_1 l_2 l_3}$ . We shall employ this mode expansion methodology in this paper, convolving observational maps with the separable basis functions and then reconstructing the

observed bispectrum  $b_{l_1 l_2 l_3}$  in an expansion using the resulting mode coefficients. Rather than looking in just a few specific directions within the large space of possible bispectra, this general mode decomposition encompasses all bispectra up to a given resolution. Our aim is to determine whether there is evidence for *any* bispectrum causing a departure from Gaussianity in the WMAP data. Of course, we can compare with previous constraints for the local and equilateral models, but an important byproduct is a set of entirely new constraints on a wide range of non-separable models.

While we believe this work represents a significant step forward, we note that this analysis is far from the last word on CMB non-Gaussianity, not least because much higher quality and higher resolution data will soon be available from Planck. We also note that we have only used WMAP5 data out to  $l=500$ , together with a pseudo-optimal analysis of the noise and masking contributions. This paper should be considered primarily as a proof of concept implementation of these methods, leaving up to an order of magnitude discovery potential available for bispectrum signals with new CMB data, let alone future galaxy and other 3D surveys where this approach can also be applied. We note that there are other recent methodologies in the literature which, in principle, can be used to extract information from the bispectrum beyond simple separable cases, including the bispectrum power approach of ref. [6]), bispectrum binning used in ref. [7] and wavelet approaches (see the review [3]).

In section II we review general results regarding primordial and angular bispectra and their optimal estimation. The eigenmode decomposition of the bispectrum that constitutes the foundation of our methodology is summarized in section III. We then show in section IV how this expansion can be used to reconstruct the full bispectrum from the data, and proceed to validate this result in section V, before directly extracting the bispectrum from WMAP data in section VI. We finally turn our attention to estimates of  $f_{\text{NL}}$  for a wide variety of shapes, including both scale invariant bispectra (section VII) and scale-dependent oscillatory bispectra (section VIII). Finally, before drawing our conclusions in section X, we discuss in section IX a possible way to use our mode expansion technique to define a model independent constraint on the total integrated bispectrum extracted from the data.

## II. CMB BISPECTRUM ESTIMATION

### A. Primordial and CMB bispectrum

Temperature anisotropies are represented using the  $a_{lm}$  coefficients of a spherical harmonic decomposition of the cosmic microwave sky,

$$\frac{\Delta T}{T}(\hat{\mathbf{n}}) = \sum_{lm} a_{lm} Y_{lm}(\hat{\mathbf{n}}), \quad (1)$$

with an (ideal) angular power spectrum  $C_l = \sum_m a_{lm} a_{l-m}$ . The CMB bispectrum is the three point correlator of the  $a_{lm}$ ,

$$B_{m_1 m_2 m_3}^{l_1 l_2 l_3} = a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3}, \quad (2)$$

where, here, we assume that the  $B_{m_1 m_2 m_3}^{l_1 l_2 l_3}$  coefficients are not an ensemble average but, instead, directly calculated using the  $a_{lm}$ 's from a high resolution map (or maps), that is, from an experiment such as WMAP or Planck. We shall assume for the moment that if there is a non-trivial bispectrum then it has arisen through a physical process which is statistically isotropic, so we can employ the angle-averaged bispectrum  $B_{l_1 l_2 l_3}$  without loss of information, that is [8],

$$\begin{aligned} B_{l_1 l_2 l_3} &= \sum_{m_i} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} B_{m_1 m_2 m_3}^{l_1 l_2 l_3} \\ &= \sum_{m_i} h_{l_1 l_2 l_3}^{-1} \mathcal{G}_{m_1 m_2 m_3}^{l_1 l_2 l_3} B_{m_1 m_2 m_3}^{l_1 l_2 l_3}, \end{aligned} \quad (3)$$

where  $h_{l_1 l_2 l_3}$  is a geometrical factor,

$$h_{l_1 l_2 l_3} = \sqrt{\frac{(2l_1 + 1)(2l_2 + 1)(2l_3 + 1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}, \quad (4)$$

and  $\mathcal{G}_{m_1 m_2 m_3}^{l_1 l_2 l_3}$  is the Gaunt integral,

$$\begin{aligned} \mathcal{G}_{m_1 m_2 m_3}^{l_1 l_2 l_3} &\equiv \int d\Omega Y_{l_1 m_1}(\hat{\mathbf{n}}) Y_{l_2 m_2}(\hat{\mathbf{n}}) Y_{l_3 m_3}(\hat{\mathbf{n}}) \\ &= h_{l_1 l_2 l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}, \end{aligned} \quad (5)$$

with the usual Wigner-3j symbol  $\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ . It is more convenient to eliminate the geometrical factors entirely and to work with the reduced bispectrum which is defined as

$$b_{l_1 l_2 l_3} = h_{l_1 l_2 l_3}^{-1} B_{l_1 l_2 l_3}. \quad (6)$$

It is important to note the relationship between the late-time CMB bispectrum  $b_{l_1 l_2 l_3}$  and the primordial bispectrum  $B_\Phi(k_1, k_2, k_3)$  from which it would arise in many models, notably inflation. The convention has been to remove a  $k^{-6}$  scaling by defining a shape function:

$$S(k_1, k_2, k_3) \equiv \frac{1}{N} (k_1 k_2 k_3)^2 B_\Phi(k_1, k_2, k_3). \quad (7)$$

The shape function (7) is particularly pertinent for scale-invariant models because their momentum dependence is restricted entirely to planes transverse to the diagonal  $\tilde{k} = \frac{1}{2}(k_1 + k_2 + k_3)$ . The CMB bispectrum induced by the primordial shape  $S$  is obtained from the convolution [9]:

$$\begin{aligned} b_{l_1 l_2 l_3} &= \left(\frac{2}{\pi}\right)^3 \int x^2 dx \int dk_1 dk_2 dk_3 S(k_1, k_2, k_3) \\ &\quad \times \Delta_{l_1}(k_1) \Delta_{l_2}(k_2) \Delta_{l_3}(k_3) j_{l_1}(k_1 x) j_{l_2}(k_2 x) j_{l_3}(k_3 x), \end{aligned} \quad (8)$$

where  $\Delta_l(k)$  is the transfer function.

The impact of the transfer functions in (8) is to impose a series of acoustic peaks on the underlying primordial shape, as illustrated for the CMB bispectrum of the constant model  $S(k_1, k_2, k_3) = 1$  in fig. 1. Here, we can observe a large primary peak when all the  $l_i \approx 220$ . In principle, the CMB bispectrum is difficult to evaluate since (8) represents a four-dimensional integral over highly oscillatory functions. However, the integral breaks down into a product of one-dimensional integrals if the shape function is separable, that is, if it can be represented in the form  $S(k_1, k_2, k_3) = X(k_1)Y(k_2)Z(k_3)$ . In the large-angle limit with  $\Delta_l(k) = j_l(\dots)$  ( $l \ll 200$ ) it is possible in some separable models to obtain analytic solutions, such as that for the constant model [5]

$$b_{l_1 l_2 l_3}^{const(la)} = \frac{\Delta_\Phi^2}{27N} \frac{1}{(2l_1 + 1)(2l_2 + 1)(2l_3 + 1)} \left[ \frac{1}{l_1 + l_2 + l_3 + 3} + \frac{1}{l_1 + l_2 + l_3} \right]. \quad (9)$$

This particular regular solution is important because we divide by it when plotting the CMB bispectrum  $b_{l_1 l_2 l_3} / b_{l_1 l_2 l_3}^{const(la)}$  throughout this paper. Normalising with the constant model (9) is analogous to multiplying the power spectrum  $C_l$ 's by  $l(l+1)$ , because it serves to remove an overall  $\ell^{-4}$  scaling for all scale-invariant bispectra, preserving the effects of the oscillating transfer functions without introducing spurious transverse momentum dependence.

## B. CMB bispectrum estimators

Now it is usually presumed that the full bispectrum for a high resolution map cannot be evaluated explicitly because of the sheer number of operations involved  $\mathcal{O}(l_{\max}^5)$ , as well as the fact that the signal

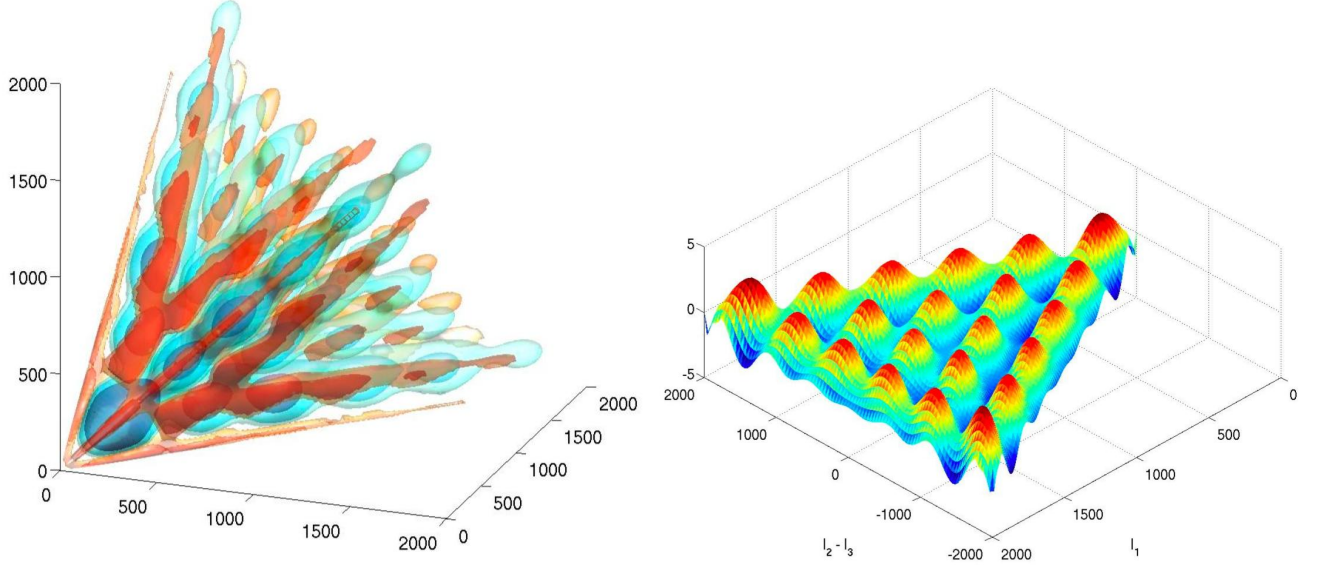


Figure 1: The reduced CMB bispectrum for the constant model  $b_{l_1 l_2 l_3}^{\text{const}}$  arising from the convolution of the primordial shape function  $S(k_1, k_2, k_3) = 1$  with transfer functions (normalised relative to the large-angle constant solution  $b_{l_1 l_2 l_3}^{\text{const}(la)}$  given in (9)). On the left, the 3D bispectrum is plotted over the allowed tetrahedral region of multipole triples (see fig. 2) using several density contours (light blue positive and magenta negative) out to  $l_i \leq 2000$ . On the right, a transverse triangular slices through the bispectrum is shown for  $l_1 + l_2 + l_3 = 4000$  (Planck resolution). Note the coherent pattern of acoustic peaks with a dominant primary peak in a broad diagonal region around  $l_1 = l_2 = l_3 = 220$ . This constant model bispectrum plotted is the analogue of the angular power spectrum  $C_l$ 's for a purely scale-invariant model.

will be too weak to measure individual multipoles with any significance. Instead, we essentially use a least squares fit to compare the bispectrum of the observed  $a_{lm}$ 's (1) with a particular (separable) theoretical bispectrum  $b_{l_1 l_2 l_3}^{\text{th}}$ ,

$$\langle a_{l_1 m_1}^{\text{th}} a_{l_2 m_2}^{\text{th}} a_{l_3 m_3}^{\text{th}} \rangle = \mathcal{G}_{m_1 m_2 m_3}^{l_1 l_2 l_3} b_{l_1 l_2 l_3}^{\text{th}}. \quad (10)$$

Here,  $b_{l_1 l_2 l_3}^{\text{th}}$  will be recovered as the expectation value from an ensemble average over  $a_{lm}^{\text{th}}$  realisations or simulations created with the given reduced bispectrum. Formally, taking into account the fact that instrument noise and masking can break rotational invariance, the result is the general optimal estimator [10–12]

$$\begin{aligned} \mathcal{E} = \frac{1}{N^2} \sum_{l_i, m_i} & \left[ \mathcal{G}_{m_1 m_2 m_3}^{l_1 l_2 l_3} b_{l_1 l_2 l_3}^{\text{th}} \left( C_{l_1 m_1, l_4 m_4}^{-1} a_{l_1 m_1} \right) \left( C_{l_2 m_2, l_5 m_5}^{-1} a_{l_2 m_2} \right) \left( C_{l_3 m_3, l_6 m_6}^{-1} a_{l_3 m_3} \right) \right. \\ & \left. - 3 \langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle C_{l_1 m_1, l_2 m_2}^{-1} C_{l_3 m_3, l_4 m_4}^{-1} a_{l_4 m_4} \right], \end{aligned} \quad (11)$$

where  $C^{-1}$  is the inverse of the covariance matrix  $C_{l_1 m_1, l_2 m_2} = \langle a_{l_1 m_1} a_{l_2 m_2} \rangle$  and  $N$  is a suitable normalisation (discussed further below). Here, we follow ref. [13, 14], by assuming a nearly diagonal covariance matrix ( $C_{l_1 m_1, l_2 m_2} \approx C_l \delta_{l_1 l_2} \delta_{m_1 - m_2}$ ) and approximating the estimator (11) as

$$\mathcal{E} = \frac{1}{N^2} \sum_{l_i, m_i} \frac{\mathcal{G}_{m_1 m_2 m_3}^{l_1 l_2 l_3} \tilde{b}_{l_1 l_2 l_3}}{\tilde{C}_{l_1} \tilde{C}_{l_2} \tilde{C}_{l_3}} (a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} - 6 C_{l_1 m_1, l_2 m_2}^{\text{sim}} a_{l_3 m_3}), \quad (12)$$

where the tilde denotes the modification of  $C_l$  and  $b_{l_1 l_2 l_3}$  to incorporate instrument beam and noise effects through

$$\tilde{C}_l = b_l^2 C_l + N_l \quad \text{and} \quad \tilde{b}_{l_1 l_2 l_3} = b_{l_1} b_{l_2} b_{l_3} b_{l_1 l_2 l_3}. \quad (13)$$

For a relatively small galactic mask (leaving a large fraction  $f_{\text{sky}}$  of the full sky), it has also been shown to be a good approximation to renormalise using

$$b_{l_1 l_2 l_3}^{\text{mask}} = f_{\text{sky}} b_{l_1 l_2 l_3} \quad \text{and} \quad C_l^{\text{mask}} = f_{\text{sky}} C_l. \quad (14)$$

(We shall assume noise, beam and mask inclusion henceforth and drop any special notation.) Here, the second linear term in (12) ensures subtraction of spurious inhomogeneous noise and masking contributions by using the covariance matrix  $C_{l_1 m_1, l_2 m_2}^{\text{sim}}$  from an ensemble average of Gaussian maps in which these effects are incorporated.

If the theoretical bispectrum  $b_{l_1 l_2 l_3}^{\text{th}}$  has the property of primordial separability then it has been noted that the summation in (12) becomes much more tractable taking only  $\mathcal{O}(l_{\text{max}}^3)$  operations [10]. Essentially this exploits the separability of the Gaunt integral (5), as well as primordial counterparts, to reduce the dimensionality of the integrals and summations involved in evaluating (12) (see ref. [1] for a more detailed discussion on this point). To date, such separability has been a property of all the primordial theories constrained observationally with most attention given to the canonical local model.

### C. $F_{\text{NL}}$ normalisation

It remains to briefly discuss the normalisation factor  $N$  in (11). In the past this has been taken on a case-by-case manner for a given theoretical bispectrum  $b_{l_1 l_2 l_3}^{\text{th}}$  to be

$$N_{\text{th}}^2 \equiv \sum_{l_i} \frac{h_{l_1 l_2 l_3}^2 b_{l_1 l_2 l_3}^{\text{th}^2}}{C_{l_1} C_{l_2} C_{l_3}}. \quad (15)$$

As we discuss below, this has yielded very model-dependent results for the measurement of the nonlinearity parameter  $f_{\text{NL}}^{\text{th}} \equiv \mathcal{E}$ . Instead, we have proposed the parameter  $F_{\text{NL}}$  which is much easier to compare between models, because it measures the integrated CMB bispectrum signal relative to that from the canonical local model with  $f_{\text{NL}} = 1$ . In this case, we define[1]

$$F_{\text{NL}}^{\text{th}} = \mathcal{E}, \quad \text{with} \quad N^2 \equiv N_{\text{loc}} N_{\text{th}}, \quad (16)$$

with  $N_{\text{th}}$  from (15) and where  $N_{\text{loc}}$  is defined for the  $f_{\text{NL}} = 1$  local model:

$$N_{\text{loc}}^2 \equiv \sum_{l_i} \frac{h_{l_1 l_2 l_3}^2 b_{l_1 l_2 l_3}^{\text{loc}(f_{\text{NL}}=1)^2}}{C_{l_1} C_{l_2} C_{l_3}}. \quad (17)$$

Of course, for the local model, the quantities are identical ( $F_{\text{NL}}^{\text{th}} = f_{\text{NL}}^{\text{loc}}$ ). However, when we quote constraints on other models we will use  $F_{\text{NL}}$ —making self-evident the comparable nature of this quantity—while also noting the  $f_{\text{NL}}^{\text{th}}$  previously used in the literature.

The problem with  $f_{\text{NL}}^{\text{th}}$  is that it derives from a somewhat arbitrary normalisation of the primordial bispectrum  $B_{\Phi}^{\text{th}}(k_1, k_2, k_3)$  which bears little relation to the observable CMB bispectrum signal. The convention has been to assume a nearly scale-invariant shape function  $S(k_1, k_2, k_3)$  and then to normalise it such that  $S^{\text{th}}(k, k, k) = 1$ , that is, at a single point; this becomes  $f_{\text{NL}} = 1$  case for the model under study. This definition ignores the behaviour away from the equilateral value  $k=k_1=k_2=k_3$ . For example,  $S$  rises from a central minimum in the local model and falls from a maximum in the equilateral model; hence, the huge disparities between their  $f_{\text{NL}}$  constraints, e.g.  $\Delta f_{\text{NL}}^{\text{equil}} \approx 7 \Delta f_{\text{NL}}^{\text{loc}}$ . This definition also does not apply to non-scaling models. The alternative to base the non-Gaussianity measure on the actually observable CMB bispectrum  $b_{l_1 l_2 l_3}^{\text{th}}$ , as above in (16), does accommodate non-scale invariant models, such as feature models. It also covers bispectra induced by late-time processes like gravitational lensing and cosmic strings. For models which are not scale-invariant it should be quoted with the observational cut-off  $l_{\text{max}}$ . The normalisation for a particular model  $F_{\text{NL}}^{\text{th}}$  can be easily forecast using the primordial  $B_{\Phi}^{\text{th}}(k_1, k_2, k_3)$  without the need for accurate CMB calculations of  $b_{l_1 l_2 l_3}^{\text{th}}$  in (8); primordial shape autocorrelators just need to be compared with the local shape as demonstrated in ref. [5].

### III. SEPARABLE MODE EXPANSIONS

When analysing the CMB bispectrum  $b_{l_1 l_2 l_3}$ , we are restricted to a tetrahedral domain of multipole triples  $\{l_1 l_2 l_3\}$  satisfying both a triangle condition and a limit given by the maximum resolution  $l_{\max}$  of the experiment. This three-dimensional domain  $\mathcal{V}_T$  of allowed multipoles is illustrated in fig. 2 and it is explicitly defined by

$$\begin{aligned} \text{Resolution:} \quad & l_1, l_2, l_3 \leq l_{\max}, \quad l_1, l_2, l_3 \in \mathbb{N}, \\ \text{Triangle condition:} \quad & l_1 \leq l_2 + l_3 \text{ for } l_1 \geq l_2, l_3, \text{ + cyclic perms.}, \\ \text{Parity condition:} \quad & l_1 + l_2 + l_3 = 2n, \quad n \in \mathbb{N}. \end{aligned} \tag{18}$$

The multipole domain is denoted a ‘tetrapyd’ because it arises from the union of a regular tetrahedron from the origin out to the plane  $l_1 + l_2 + l_3 \leq 2l_{\max}$  and a triangular pyramid constructed from the corner of the cube taking in the remaining multipole values out to  $l_i \leq l_{\max}$ . Summed bispectrum expressions such as (15) indicate that we must define a weight function  $w_{l_1 l_2 l_3}$  on the tetrapyd domain in terms of the geometrical factor  $h_{l_1 l_2 l_3}$ , that is,

$$w_{l_1 l_2 l_3} = h_{l_1 l_2 l_3}^2. \tag{19}$$

This is a nearly constant function on cross sections defined by  $l_1 + l_2 + l_3 = \text{const}$ , except very near the tetrahedral boundaries where it is still bounded, and a useful and accurate continuum limit  $w(l_1, l_2, l_3)$  is given in [1]. In order to eliminate an  $l^{-1/2}$  scaling in the bispectrum estimator functions, we usually exploit the freedom to divide by a separable function and to employ instead the weight

$$w_s(l_1, l_2, l_3) = \frac{w_{l_1 l_2 l_3}}{v_{l_1}^2 v_{l_2}^2 v_{l_3}^2}, \quad \text{where } v_l = (2l + 1)^{1/6}. \tag{20}$$

We can then define an inner product of two functions  $f(l_1, l_2, l_3)$ ,  $g(l_1, l_2, l_3)$  on the tetrapyd domain (18) through

$$\langle f, g \rangle \equiv \sum_{l_1, l_2, l_3 \in \mathcal{V}_T} w_s(l_1, l_2, l_3) f(l_1, l_2, l_3) g(l_1, l_2, l_3). \tag{21}$$

Given that calculations generally deal with smooth functions  $f, g, w, v$ , we can use a variety of schemes to speed up this summation (effectively an integration).

Our goal is to represent the observed CMB bispectrum estimator functions, such as those in (12) and (15), on the multipole domain (18) using a separable mode expansion,

$$\frac{v_{l_1} v_{l_2} v_{l_3}}{\sqrt{C_{l_1} C_{l_2} C_{l_3}}} b_{l_1 l_2 l_3} = \sum_n \bar{\alpha}_n^\omega \bar{\mathcal{Q}}_n(l_1, l_2, l_3), \tag{22}$$

where the  $\bar{\mathcal{Q}}_n$  are basis functions constructed from symmetrised polynomial products

$$\begin{aligned} \bar{\mathcal{Q}}_n(l_1, l_2, l_3) &= \frac{1}{6} [\bar{q}_p(l_1) \bar{q}_r(l_2) \bar{q}_s(l_3) + \bar{q}_r(l_1) \bar{q}_p(l_2) \bar{q}_s(l_3) + \text{cyclic perms in } prs] \\ &\equiv \bar{q}_{\{pqr\}} \quad \text{with } n \leftrightarrow \{prs\}, \end{aligned} \tag{23}$$

with the  $\bar{q}_p(l)$  defined below. Here, the six permutations of the polynomial products which we denote as  $\{prs\}$  reflect the underlying symmetries of the bispectrum  $b_{l_1 l_2 l_3}$ . For convenience, we define a one-to-one mapping  $n \leftrightarrow \{prs\}$  ordering the permuted triple indices into a single list labelled by  $n \in \mathbb{N}$ . Alternative ‘slicing’ and ‘distance’ orderings were presented in ref. [1], but the results presented here are robust to this change. However, we shall quote explicit coefficients  $\beta_n^\omega$  resulting from distance ordering (i.e.  $n(l_1, l_2, l_3) < n'(l'_1, l'_2, l'_3)$  implies  $l_1^2 + l_2^2 + l_3^2 \leq l'_1{}^2 + l'_2{}^2 + l'_3{}^2$  and in the instance of two modes being equidistant the one with most equal  $l_i$  takes precedence).

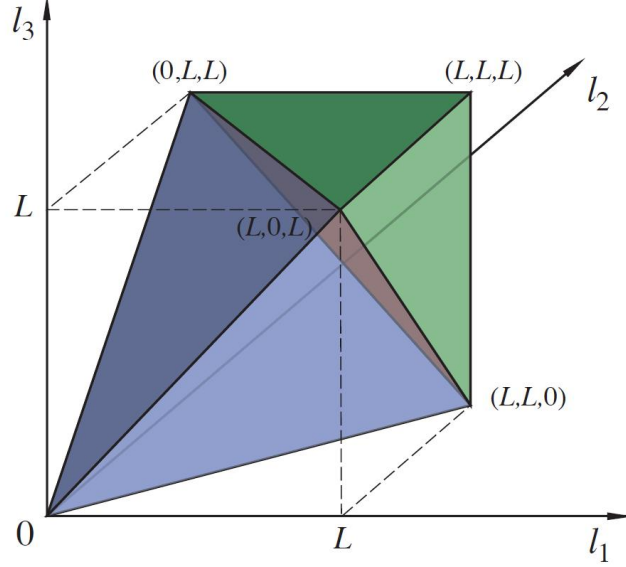


Figure 2: Observational domain (18) for the CMB bispectrum  $b_{l_1 l_2 l_3}$ . Allowed multipole values  $(l_1, l_2, l_3)$  lie inside the shaded ‘tetrapyd’ region, satisfying both the triangle condition and  $l < L \equiv l_{\max}$ .

We choose to define the tetrahedral  $\bar{q}_p(l)$  polynomials analogously to Legendre polynomials  $P_n$  by requiring them to be self-orthogonal with respect to the inner product (21),

$$\langle \bar{q}_p(l_1), \bar{q}_r(l_1) \rangle = \delta_{pr}, \quad (24)$$

with the first few polynomials given by  $\bar{q}_0 = 0.074$ ,  $\bar{q}_1 = 0.30(-0.61 + l)$ ,  $\bar{q}_2 = 1.2(0.26 - 1.1l + l^2)$  etc. More precise expressions and generating functions are given in ref. [1]. As products, the  $q_p$  only confer partial orthogonality on the 3D basis functions  $\bar{Q}_n$ , but their use is vital for other reasons, given their bounded and near scale-invariant behaviour.

While the product basis functions  $\bar{Q}_n$  are independent and separable, they are not orthogonal in general

$$\langle \bar{Q}_n, \bar{Q}_p \rangle \equiv \gamma_{np} \neq \delta_{np}, \quad (25)$$

so it is very useful to construct a related set of orthonormal mode functions  $\bar{\mathcal{R}}_n$  using Gram-Schmidt orthogonalisation such that

$$\langle \bar{\mathcal{R}}_n, \bar{\mathcal{R}}_p \rangle = \delta_{np}. \quad (26)$$

Working up to a given order  $N$ , the two sets of mode functions are related through

$$\mathcal{R}_n = \sum_{p=0}^n \lambda_{np} \mathcal{Q}_p \quad \text{for } n, p \leq N, \quad (27)$$

where  $\lambda_{mp}$  is a lower triangular matrix with

$$(\lambda^{-1})_{np}^\top = \langle \mathcal{Q}_n, \mathcal{R}_p \rangle \quad \text{and} \quad (\gamma^{-1})_{np} = \sum_r^N (\lambda^\top)_{nr} \lambda_{rp}. \quad (28)$$

Knowing  $\lambda_{np}$  allows us to easily systematically evaluate the expansion coefficients in (22) directly from the inner product

$$\bar{\alpha}_n^\mathcal{R} = \left\langle \bar{\mathcal{R}}_n, \frac{v_{l_1} v_{l_2} v_{l_3}}{\sqrt{C_{l_1} C_{l_2} C_{l_3}}} b_{l_1 l_2 l_3} \right\rangle, \quad \text{yielding} \quad \bar{\alpha}_n^\mathcal{Q} = \sum_{p=0}^N (\lambda^\top)_{np} \bar{\alpha}_p^\mathcal{R}. \quad (29)$$

Indeed, it is more convenient to present our final bispectrum results in the orthonormal  $\bar{\mathcal{R}}_n$  basis,

$$\frac{v_{l_1} v_{l_2} v_{l_3}}{\sqrt{C_{l_1} C_{l_2} C_{l_3}}} b_{l_1 l_2 l_3} = \sum_n \bar{\alpha}_n^{\mathcal{R}} \bar{\mathcal{R}}_n \quad (30)$$

because their orthonormality (26) implies a version of Parseval's theorem. Here, we note that the expansion (30) presumes a spectrum normalised as in (16) to have  $F_{\text{NL}} = 1$ , that is, with  $N$  such that  $\sum_n \bar{\alpha}_n^{\mathcal{R}^2} = N^2$  in the estimator (12).

To summarise, the  $\bar{\mathcal{Q}}_n(l_1, l_2, l_3)$ 's are independent separable basis functions built out of the permutations of simple products of the polynomials  $\bar{q}_p(l)$ , which are well-behaved and bounded over the tetrapyd. The  $\bar{\mathcal{Q}}_n$ 's in their easily separable form are employed directly in the bispectrum estimator. However, it is more straightforward to present results and to use the inner product (21) with the transformed  $\bar{\mathcal{R}}_n$  eigenmodes because they are orthonormal; a simple matrix expression (29) relates the expansion coefficients  $\bar{\alpha}_n^{\mathcal{Q}}$  and  $\bar{\alpha}_n^{\mathcal{R}}$  using the two sets of basis functions.

#### IV. RECONSTRUCTING THE CMB BISPECTRUM

Now consider the implications of substituting the mode expansion (22) into the estimator (12), while exploiting the separability of the Gaunt integral (5),

$$\begin{aligned} \mathcal{E} &= \frac{1}{N^2} \sum_{l_i, m_i} \sum_{n \leftrightarrow prs} \bar{\alpha}_n^{\mathcal{Q}} \bar{q}_{\{p} \bar{q}_r \bar{q}_{s\}} \int d^2 \hat{\mathbf{n}} \frac{Y_{l_2 m_2}(\hat{\mathbf{n}}) Y_{l_1 m_1}(\hat{\mathbf{n}}) Y_{l_3 m_3}(\hat{\mathbf{n}})}{v_{l_1} v_{l_2} v_{l_3} \sqrt{C_{l_1} C_{l_2} C_{l_3}}} [a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} - 6 \langle a_{l_1 m_1} a_{l_2 m_2} \rangle a_{l_3 m_3}] \\ &= \frac{1}{N^2} \sum_{n \leftrightarrow prs} \bar{\alpha}_n^{\mathcal{Q}} \int d^2 \hat{\mathbf{n}} \left[ \left( \sum_{l_1, m_1} \bar{q}_{\{p} \frac{a_{l_1 m_1} Y_{l_1 m_1}}{v_{l_1} \sqrt{C_{l_1}}} \right) \left( \sum_{l_2, m_2} \bar{q}_r \frac{a_{l_2 m_2} Y_{l_2 m_2}}{v_{l_2} \sqrt{C_{l_2}}} \right) \left( \sum_{l_3, m_3} \bar{q}_{s\} \frac{a_{l_3 m_3} Y_{l_3 m_3}}{v_{l_3} \sqrt{C_{l_3}}} \right) \right. \\ &\quad \left. - 6 \left\langle \left( \sum_{l_1, m_1} \bar{q}_{\{p} \frac{a_{l_1 m_1} Y_{l_1 m_1}}{v_{l_1} \sqrt{C_{l_1}}} \right) \left( \sum_{l_2, m_2} \bar{q}_r \frac{a_{l_2 m_2} Y_{l_2 m_2}}{v_{l_2} \sqrt{C_{l_2}}} \right) \right\rangle \left( \sum_{l_3, m_3} \bar{q}_{s\} \frac{a_{l_3 m_3} Y_{l_3 m_3}}{v_{l_3} \sqrt{C_{l_3}}} \right) \right] \quad (31) \end{aligned}$$

$$= \frac{1}{N^2} \sum_{n \leftrightarrow prs} \bar{\alpha}_n^{\mathcal{Q}} \int d^2 \hat{\mathbf{n}} \left[ \bar{M}_{\{p}(\hat{\mathbf{n}}) \bar{M}_r(\hat{\mathbf{n}}) \bar{M}_{s\}}(\hat{\mathbf{n}}) - 6 \langle \bar{M}_{\{p}^G(\hat{\mathbf{n}}) \bar{M}_r^G(\hat{\mathbf{n}}) \rangle \bar{M}_{s\}}(\hat{\mathbf{n}}) \right]. \quad (32)$$

Here, the  $\bar{M}_p(\hat{\mathbf{n}})$  represent versions of the original CMB map filtered with the polynomial  $\bar{q}_p$  with the separated weight function  $(v_l \sqrt{C_l})^{-1}$ , that is,

$$\bar{M}_p(\hat{\mathbf{n}}) = \sum_{lm} \bar{q}_p(l) \frac{a_{lm}}{v_l \sqrt{C_l}} Y_{lm}(\hat{\mathbf{n}}). \quad (33)$$

The maps  $\bar{M}_p^G(\hat{\mathbf{n}})$  incorporate the same mask and a realistic model of the inhomogeneous instrument noise; a large ensemble of these maps, calculated from Gaussian simulations, are used in the averaged linear term in the estimator (31), allowing for the subtraction of these important effects. Defining the integral over these convolved product maps as cubic and linear terms respectively

$$\begin{aligned} \bar{\beta}_n^{\mathcal{Q}^{\text{cub}}} &= \int d^2 \hat{\mathbf{n}} \bar{M}_{\{p}(\hat{\mathbf{n}}) \bar{M}_r(\hat{\mathbf{n}}) \bar{M}_{s\}}(\hat{\mathbf{n}}), \\ \bar{\beta}_n^{\mathcal{Q}^{\text{lin}}} &= \int d^2 \hat{\mathbf{n}} \langle \bar{M}_{\{p}^G(\hat{\mathbf{n}}) \bar{M}_r^G(\hat{\mathbf{n}}) \rangle \bar{M}_{s\}}(\hat{\mathbf{n}}), \end{aligned} \quad (34)$$

the estimator (12) reduces to a simple sum over the mode coefficients

$$\mathcal{E} = \frac{1}{N^2} \sum_n \bar{\alpha}_n^{\mathcal{Q}} \bar{\beta}_n^{\mathcal{Q}}, \quad (35)$$



where  $\bar{\beta}_n^\varnothing \equiv \bar{\beta}_n^{\varnothing\text{cub}} - \bar{\beta}_n^{\varnothing\text{lin}}$ .

The estimator sum (35) is straightforward to evaluate, provided the theoretical model coefficients  $\bar{\alpha}_n^\varnothing$  are known. It has been separated into a product of three sums over the observational maps (31), followed by a single integral over all directions (34). The actual operations entailed in the estimator sum are only  $\mathcal{O}(l^2)$ , so these late-time methods are extremely rapid for direct data analysis and for obtaining variances from map simulations. However, we note that the preparatory ‘one-off’ calculations setting up the orthonormal eigenmodes and theoretical CMB bispectra are of order  $\mathcal{O}(l^3)$ . We emphasise that the utility of this approach depends on a fairly rapidly convergent expansion for the theoretical bispectrum under study (as indicated for almost all models studied to date [5]) and the fact that we have constructed a *complete* set of orthonormal eigenmodes on the observed multipole domain (18).

There is potentially much more information in the observed  $\bar{\beta}_n^\varnothing$  coefficients than just the estimator sum (35) which only yields  $f_{\text{NL}}$  for a given theoretical model. Following the steps above in (31), it is easy to show (see Appendix) that the expectation value for  $\bar{\beta}_n^\varnothing$  for an ensemble of maps with a given CMB bispectrum (expanded in modes  $\bar{\alpha}_n^\mathcal{R}$  with amplitude  $F_{\text{NL}}$ ) is

$$\langle \bar{\beta}_n^\varnothing \rangle = \sum_p F_{\text{NL}} \bar{\alpha}_n^\varnothing \langle \bar{\mathcal{Q}}_n, \bar{\mathcal{Q}}_p \rangle = F_{\text{NL}} \sum_p \bar{\alpha}_n^\varnothing \gamma_{np}, \quad (36)$$

so that the averaged estimator (35) becomes

$$\langle \mathcal{E} \rangle = \frac{1}{N^2} F_{\text{NL}} \sum_n \sum_p \bar{\alpha}_n^\varnothing \gamma_{np} \bar{\alpha}_p^\varnothing = \frac{1}{N^2} F_{\text{NL}} \sum_n \bar{\alpha}_n^{\mathcal{R}2} = F_{\text{NL}}, \quad (37)$$

where we have used (29) in transforming to the  $\bar{\mathcal{R}}_n$  basis. (Here we note again that in this basis  $N^2 = \sum_n \bar{\alpha}_n^{\mathcal{R}2}$ .) Equivalently, then, in the orthonormal frame we have the simple result

$$\langle \bar{\beta}_n^\mathcal{R} \rangle = F_{\text{NL}} \bar{\alpha}_n^\mathcal{R}, \quad (38)$$

that is, we expect the best fit  $\bar{\beta}_n^\mathcal{R}$  coefficients for a particular realization to be the  $\bar{\alpha}_n^\mathcal{R}$ ’s themselves (given a sufficiently large signal). Assuming that we can extract the  $\bar{\beta}_n^\mathcal{R}$  coefficients with sufficient significance from a particular experiment, this means that we can directly reconstruct the CMB bispectrum using the expansion (30).

## V. VALIDATION USING SIMULATED MAPS

In our previous paper [1], we have validated this mode expansion approach with simulated maps in a WMAP realistic context, accurately obtaining the input  $f_{\text{NL}}$  as well as a fairly good recovery of the actual bispectrum coefficients  $\bar{\alpha}_n^\mathcal{R}$ . In that case, we chose to use the equilateral model because the noise analysis requirements were not as stringent. Here, however, we wish to undertake a comprehensive reconstruction of the bispectrum, so we have ensured an accurate and robust implementation of the full noise and mask subtraction, incorporated in the linear correction term to the bispectrum estimator (12).

As well as the equilateral model, the other canonical test model we have used to validate our methods is the local model, characterised by the shape function (7):

$$eq : localSS^{\text{local}}(k_1, k_2, k_3) = \frac{1}{3} \left( \frac{k_1^2}{k_2 k_3} + \frac{k_2^2}{k_1 k_3} + \frac{k_3^2}{k_1 k_2} \right). \quad (39)$$

The local model (39) is dominated by signal from squeezed triangles, e.g.  $k_1 \ll k_2, k_3$ . This behaviour is strongly reflected in the CMB bispectrum  $b_{l_1 l_2 l_3}$  where the impact of the transfer functions  $\Delta_l(k)$  is to impose a series of acoustic peaks on the underlying local shape. This dominant signal along the edges of the tetrapyd for the local model provides a rigorous test for subtracting the inhomogeneous noise and masking effects. This is because these effects also exhibit an overall local shape (as we will discuss in a future publication [15]).

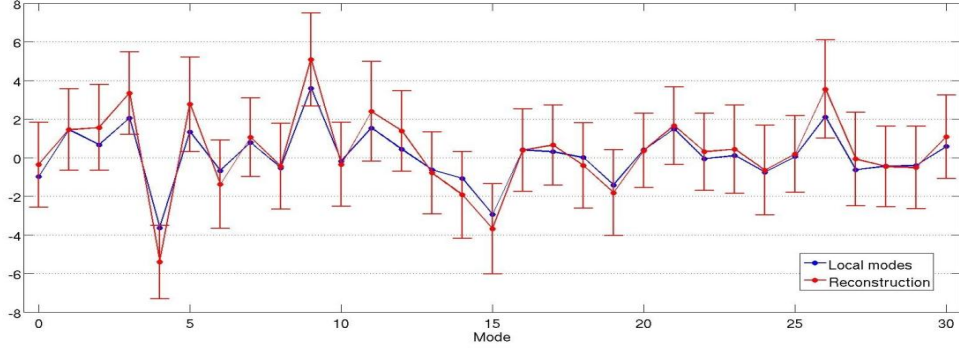


Figure 3: Recovered spectral coefficients  $\bar{\beta}_n^{\mathcal{R}}$  (22) from a single map simulation for a local model with  $f_{\text{NL}} = 100$ . The original  $\bar{\alpha}_n^{\mathcal{R}}$  decomposition coefficients for the local model are shown for comparison (blue). The  $\bar{\beta}_n^{\mathcal{R}}$  coefficients were recovered in a WMAP-realistic context using the KQ75 mask with inhomogeneous noise added. Error bars ( $1\sigma$ ) are also shown for each mode as estimated from 1000 Gaussian maps; note that this variance is roughly constant for all modes.

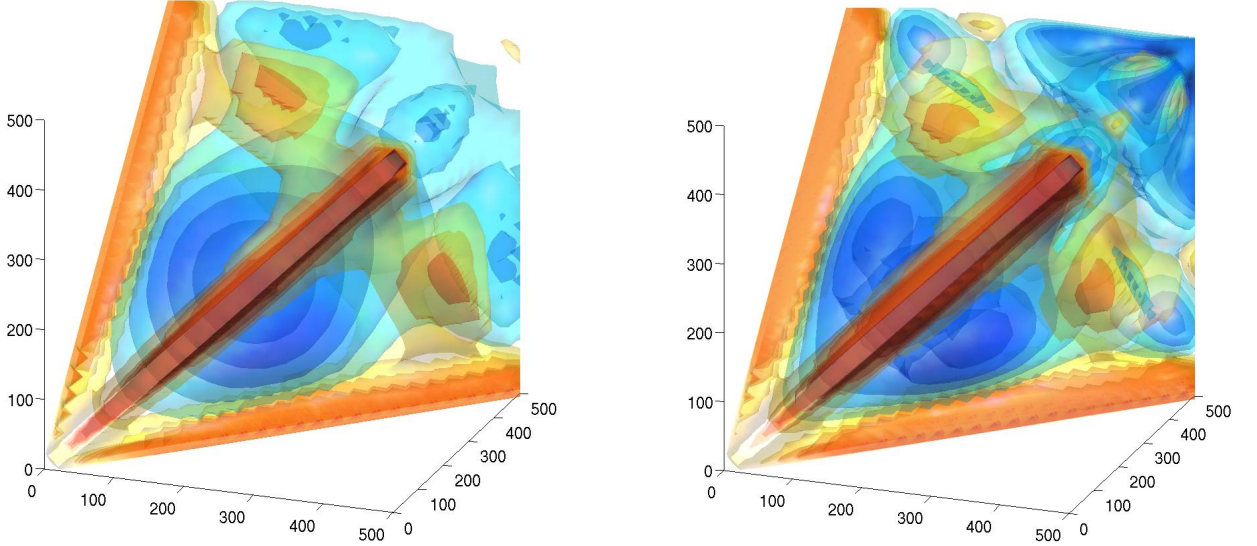


Figure 4: Recovered 3D bispectrum using the mode decomposition method (30) from the  $f_{\text{NL}} = 100$  local model map simulation used in fig. 3. The left panel represents the original theoretical bispectrum used to construct map realisations. The right panel represents the recovery in a WMAP-realistic context using the spectrum shown in fig. 3. The main features of the bispectrum are evident, including the primary acoustic peak and the high signal in the squeezed states along the tetrahedron edges.

The  $\bar{\alpha}_n^{\mathcal{R}}$  coefficients for the local bispectrum  $b_{l_1 l_2 l_3}^{\text{local}}$  are illustrated in fig. 3, having been found using the robust CMB bispectrum calculation methods described in ref. [5] (which achieve better than 99% correlation with the exact results). With only  $n_{\text{max}} = 31$   $\bar{\mathcal{Q}}_n$  modes it was possible to achieve a 94% correlation between the partial sum (22) and the original  $b_{l_1 l_2 l_3}$  for all the models studied (usually greater than 98%). These coefficients were then used to create ensembles of local maps with  $f_{\text{NL}} = 100$  (resolution  $l_{\text{max}} = 500$ ) using the separable mode map simulation methods described in ref. [1]. As a further test independent of modal map-making, the local map simulations from ref. [16] were also used to verify the main conclusions. In addition, inhomogeneous noise obtained by coadding WMAP V and W channels was included together with the use of a KQ75 sky mask, just as in the original non-Gaussian analysis of WMAP5 data.

The efficacy of the modal estimator (35) in recovering the correct  $f_{\text{NL}} = 100$  from the local maps is illustrated in fig. 3, with the ensemble average  $f_{\text{NL}} = 93 \pm 32.5$  found to be in good agreement. We expect

the value to be slightly low as we are 94% correlated (the maps have been created using an exact method rather than from the modes to ensure robustness). This confirmed the results for the equilateral model studied in ref. [1], where it was also shown that Gaussian maps give unbiased results around  $f_{\text{NL}} = 0$ . The recovery of the local bispectrum mode coefficients (34) also proved to be remarkably efficient as illustrated in fig. 3 for a typical spectrum obtained from a single map realization. The dominant local modes are clearly identifiable above the noise (for a signal of this  $3\sigma$  significance), with the typical variance obtained from Gaussian maps also shown. The three-dimensional reconstruction for the local model bispectrum is illustrated on the tetrapyd domain in fig. 4 and is comparable with the original bispectrum. The dominant features are recovered, including the primary acoustic peak at  $l_1 \approx l_2 \approx l_3 \approx 200$  and the strong signal for the squeezed triangles along the edges of the tetrahedron where one of  $l_i \approx 0$ . Comparing with results for the equilateral model in ref. [1], it is clear that for a measurement of  $3\sigma$  significance or more, we should be able to distinguish between families of models which are weakly correlated, such as local and equilateral.

## VI. THE WMAP BISPECTRUM

We now move on to apply the mode decomposition techniques described and validated in the previous sections to the analysis of WMAP5 data. Our aim, first, will be to estimate  $f_{\text{NL}}$  arising from different primordial shapes, some as yet unconstrained in the literature (such as the feature models of section VIII and the flattened models of section VII D). Secondly, we aim to provide a full reconstruction of the bispectrum from the data, using the same pipeline shown to recover local and equilateral bispectra from simulated data. The main emphasis of this work is obtaining fast and accurate convergence for many different shapes, rather than a fully optimised estimation. The analysis presented here is intended as a proof-of-concept for late time modal estimators of non-Gaussianity, gleanig valuable new information from WMAP rather achieving a maximal extraction. For this reason our study has a number of limitations, which we enumerate here. We do not implement full inverse covariance weighting in the estimator as in (11) [12], but we adopt the pseudo-optimal weighting scheme used by the WMAP team for the WMAP 5-year analysis [13]; we use multipoles up to  $\ell_{\text{max}} = 500$ , rather than 1000, since the pseudo-optimal  $f_{\text{NL}}$  error bars tend to saturate above that threshold; finally, we work with WMAP 5-year instead of WMAP 7-year data. The reason for not using the latest available dataset is not only that this work started well before the 7-year WMAP data release, but is also due to the fact that WMAP 5-year data was originally studied with a pseudo-optimal weighting approach, thus making a comparison between our results more straightforward. The present work represents the initial implementation of this general approach to analysing non-Gaussianity, rather than its

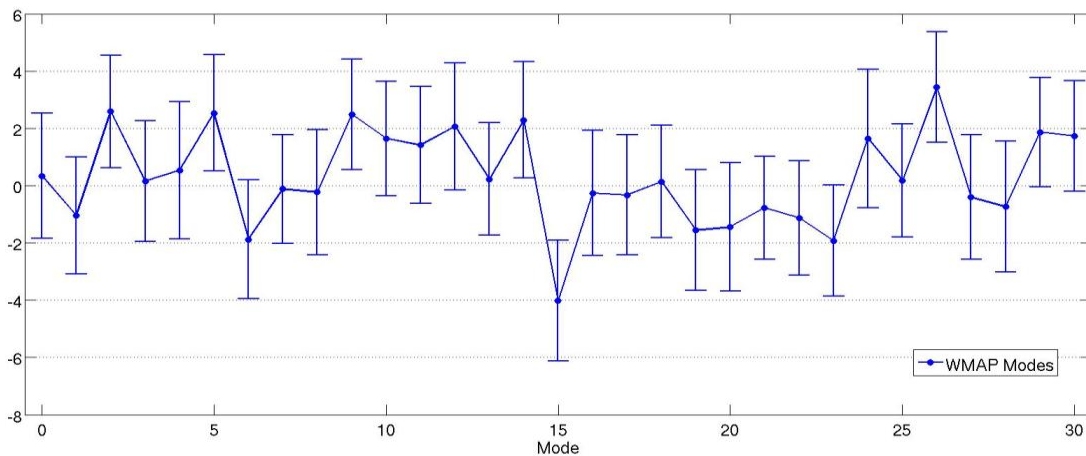


Figure 5: Recovered mode coefficients  $\bar{\beta}_n^{\mathcal{R}}$  (22) from the WMAP5 coadded V and W maps. Error bars ( $1\sigma$ ) are also shown for each mode as estimated from 1000 Gaussian map simulations in WMAP-realistic context.

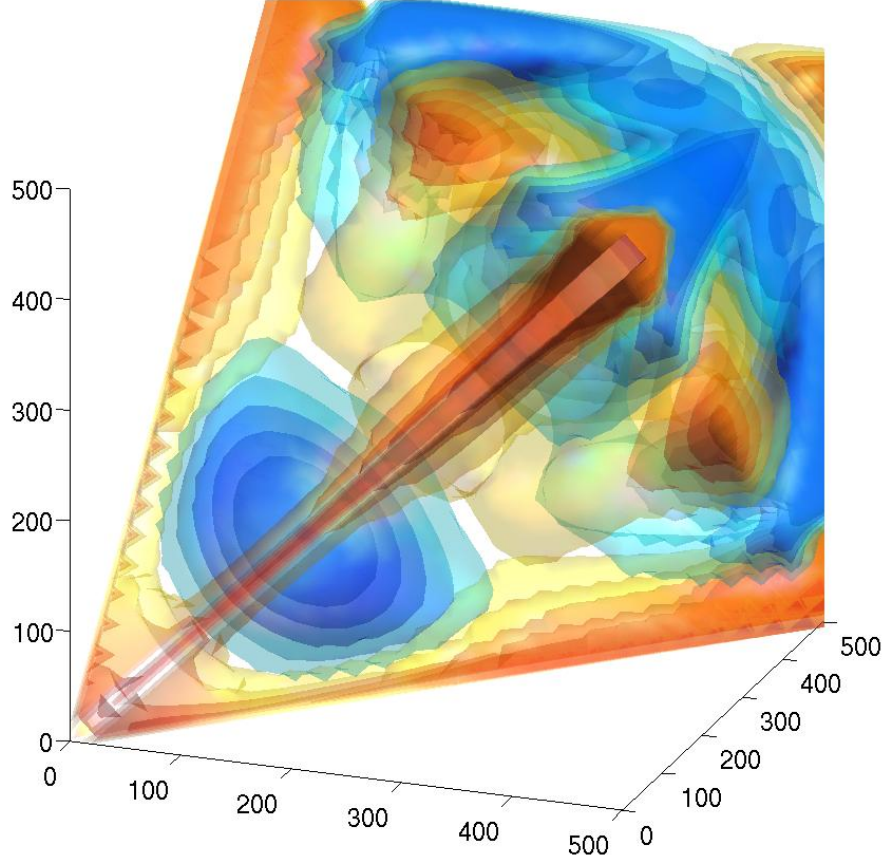


Figure 6: Recovered 3D bispectrum from WMAP5 data showing the result using the reconstructed mode coefficients  $\bar{\beta}_n^{\mathcal{R}}$  shown in fig. 5 with the partial sum (30). Several density contours (light blue positive and magenta negative) are shown out to  $l_i \leq 500$ .

completion.

After coadding the V and W band data (with the same weights as in the WMAP5 analysis), our first step was to extract the  $\bar{\beta}_n^{\mathcal{Q}}$  mode coefficients from the data, following the procedure summarized eqn (33) and (34). In our analysis we chose to compute the first  $n = 31$  modes in (35) because this proved sufficient to describe almost all theoretical CMB bispectra on the observational domain  $l_{\max} = 500$ . The resulting estimates will be shown in the following sections. As pointed out in (38), by rotating our recovered  $\bar{\beta}_n^{\mathcal{Q}}$  into the orthonormal frame we obtain the best-fit estimate of the actual bispectrum coefficients  $\bar{\alpha}_n^{\mathcal{R}}$ . The mode coefficients obtained from the WMAP5 data  $\bar{\beta}_n^{\mathcal{R}}$  in this orthonormal frame are plotted in fig. 5. The variance is estimated from 1000 Gaussian map simulations, using the pipeline repetitively in the same WMAP-realistic context.

The mode coefficient extraction from the WMAP5 data was straightforward with both the cubic and linear terms contributing significantly to the final result. The late-time estimator (12) is sensitive to all forms of non-Gaussianity, in contrast to the two or three separable (and oscillating) modes previously extracted from the data using primordial estimators. Despite this increased sensitivity, in principle, making the method more susceptible to foreground contamination, our results do not appear to have been significantly affected after subtraction by the linear term. This has been investigated through extensive testing, including increasing mask size, and we will discuss these issues at much greater length in a companion paper [15], characterising the mask, noise and other contributions. It is interesting to note here, however, that the

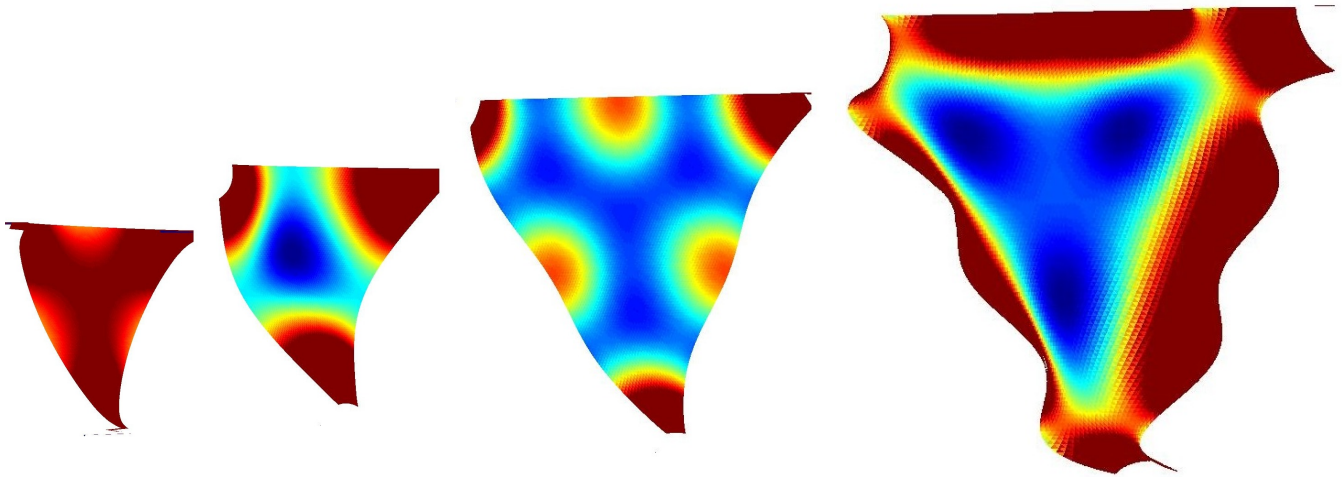


Figure 7: Recovered 3D bispectrum from WMAP5 data showing slices through the data at  $\tilde{l} \equiv l_1 + l_2 + l_3 = \text{const.}$ . Slices shown are  $\tilde{l} = 250, 500, 750, 1000$ , using the same colour scale as fig. 6.

mode decompositions can also be used to characterise anisotropic contributions, such as the inhomogeneous noise (and possibly other contaminants). We will show quantitatively how the action of the linear term essentially projects out these spurious bispectrum directions from the cubic term, with the local shape being the most affected (as noted originally in ref. [11]).

The extracted mode coefficients  $\bar{\beta}_n^\kappa$  from fig. 5 can be used to reconstruct the full 3D WMAP bispectrum using (30). The result of this partial sum is shown in fig. 6, together with a series of transverse slices through the bispectrum shown in fig. 7. Visually the WMAP bispectrum bears a qualitative similarity to the local CMB bispectrum already used for pipeline validation, illustrated in fig. 4. As we shall discuss, there appears to be some local signal emerging from the WMAP data, but the periodicity of the other features does not match well with scale-invariant primordial models. Nevertheless, the orthonormal mode coefficients  $\bar{\beta}_n^\kappa$  plotted in fig. 5 do not individually show significant deviations away from Gaussianity, presuming the accuracy of the nearly constant mode variances which are also plotted. We note at the outset, therefore, that the WMAP bispectrum shown in fig. 6 is likely to be the result of cosmic variance (perhaps with some residual local signal left-over from the noise/mask subtraction and/or other contamination). As well as constraining specific theoretical models, we shall test the assumption of Gaussianity more generally in section IX by considering a measure of the total integrated bispectrum obtained from the squared coefficients  $\bar{\beta}_n^{\kappa^2}$ . In the near future, using the full WMAP7 data set and smaller variances we will expand the scope of our mode exploration, including principal component analysis and other statistical approaches [15].

Before obtaining specific new constraints, we emphasise again that the extraction of the mode coefficients  $\bar{\beta}_n^\kappa$  provides a completely model-independent assessment of the three-point correlation function. The approach provides far more information than that contained in a simple  $f_{\text{NL}}$  amplitude parameter extraction for particular models. Although obvious deviations from Gaussianity are not apparent from this limited WMAP5 analysis (i.e. pseudo-optimal error bars and  $l_{\text{max}} = 500$ ), there remains considerable potential with new data sets. For Planck, the sensitivity to primordial non-Gaussianity will improve by up to an order of magnitude and so the error bars in fig. 5 will shrink dramatically. The prospects for detection of a large NG signal remain completely open.

## VII. CONSTRAINTS ON NEARLY SCALE-INVARIANT MODELS

Constraints on the bispectrum to date have been for scale-invariant models of separable form, primarily on the local and equilateral models, discussed previously. There has been significant evolution over time for these constraints as both the CMB data and the estimation methodology have improved. However,



as table I illustrates (taken from ref. [3]), there is no compelling and confirmed evidence for a significant non-Gaussian signal at this stage. Our purpose in this section is to apply our more general mode expansion estimator (35) with our WMAP analysis to obtain constraints on a much wider set of scale-invariant models. This method can be applied to any model for which there is good convergence with the given  $n_{\text{max}}$  modes.

	Local	Equilateral
<b>Pure cubic</b>	$-58 < f_{\text{NL}} < 134$ W1, Komatsu et al 2003 [17] $-54 < f_{\text{NL}} < 114$ W3, Spergel et al 2007 [18]	$-366 < f_{\text{NL}} < 238$ W1, Creminelli et al 2006 [11] $-256 < f_{\text{NL}} < 332$ W3, Creminelli et al 2006 [11]
<b>Pseudo-optimal</b>	$-27 < f_{\text{NL}} < 121$ W1, Creminelli et al, 2006 [11] $-36 < f_{\text{NL}} < 100$ W3, Creminelli et al 2006 [11] $27 < f_{\text{NL}} < 147$ W3, Yadav Wandelt 2008 [14] $9 < f_{\text{NL}} < 129$ W3, Smith et al 2009 [12] $-9 < f_{\text{NL}} < 111$ W5, Komatsu et al 2009 [13]	$-151 < f_{\text{NL}} < 253$ W5, Komatsu et al 2009 [13]
<b>Optimal</b>	$12 < f_{\text{NL}} < 104$ W3, Smith et al 2009 [12] $-4 < f_{\text{NL}} < 80$ W5, Smith et al 2009 [12] $-10 < f_{\text{NL}} < 74$ W7, Komatsu et al 2010 [2]	$-125 < f_{\text{NL}} < 435$ W5, Smith et al 2009 [12] $-254 < f_{\text{NL}} < 306$ W7, Komatsu et al 2010 [2]

Table I: Constraints on  $f_{\text{NL}}^{\text{local}}, f_{\text{NL}}^{\text{equil.}}$ , obtained by different groups on the one-year (W1), three-year (W3), five-year (W5), and seven-year (W7) WMAP data releases. The estimators employed are the pseudo-optimal (12), the cubic (the same without the linear noise term), and the optimal with full-covariance weighting (11). All results were in the context of a primordial estimator using separable functions to describe the specific model, unlike the general late-time estimator employed here. For further details about the estimator methods employed and the significant evolution of these results over time, please refer to the review [3].

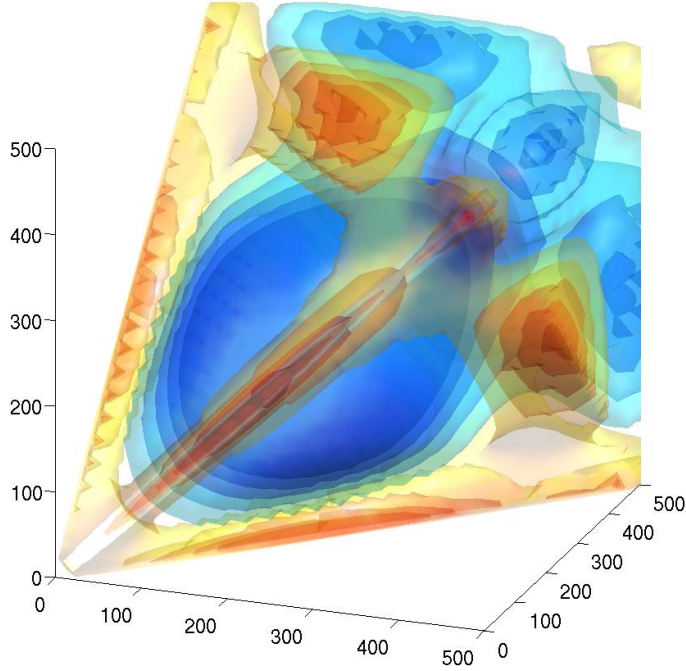


Figure 8: Predicted 3D bispectrum for the constant model up to  $l_i \leq 500$ . The same thresholds are employed as those shown in the WMAP reconstructions in fig. 4 (after an overall rescaling).

### A. Constant model

The constant model  $S(k_1, k_2, k_3) = 1$  is the simplest possible primordial shape with triangles of every configuration contributing equally, resulting in a CMB bispectrum  $b_{l_1 l_2 l_3}$  with features due entirely to the transfer functions (as we observed for the acoustic peaks shown in fig. 1). The constant model was motivated initially by its simplicity with the large-angle analytic solution (9) for the CMB bispectrum [5]. However, the constant shape does have other more explicit physical motivation, such as generation during a slowly turning trajectory in multifield inflation, denoted quasi-single field inflation [19]. For nearly scale-invariant models, the central values for the bispectrum,  $b_{lll}$ , all have roughly the same profile but with different normalisations. The oscillatory properties of the transfer functions create acoustic peaks located at triple combinations involving the following multipole values,  $l \approx 200, 500, 800, \dots$ . To observe the key differences between scale invariant models we must study the bispectrum in the plane orthogonal to the  $(l, l, l)$ -direction, that is, the directions reflecting changes in the primordial shape functions.

For the multipole range  $l_{\max} < 500$  relevant to the present analysis, we have plotted the 3D bispectrum in fig. 8. Here, the dominant feature is the primary acoustic peak stretched along the diagonal of the tetrapyd, peaking at  $l = l_1 = l_2 = l_3 = 220$  and elongated like an extended balloon from  $l \approx 100$  to  $l \approx 450$ . Evidence for this primary peak would indicate the presence of a primordial and scale-invariant non-Gaussian signal, as emphasised in ref. [5] and investigated quantitatively for the local model in ref. [7]. Observing the reconstructed WMAP bispectrum shown in fig. 6 there is a central fluctuation at  $l \approx 140$  but it does not extend to larger  $l$  as would be expected; see the  $l_1 + l_2 + l_3 = 750$  slice in fig. 7 (right) corresponding to  $l \approx 250$  where the (apparent) WMAP peak has disappeared. If this measured 3D WMAP bispectrum were considered to have any statistical significance then it would mitigate against a scale-invariant model, motivating the discussion in section VIII.

A comparison of the mode coefficients  $\bar{\alpha}_n^{\mathcal{R}\text{const}}$  from the constant model CMB bispectrum shown in fig. 9 indicates little obvious correlation with the WMAP coefficients  $\bar{\beta}_n^{\mathcal{R}\text{wmap}}$  (also plotted). Note that the constant model mode coefficients are large for the constant offset  $n = 0$  and for  $n = 3, 4, 5$  reflecting the periodicity of the acoustic peak structure (for  $l_{\max} = 500$ ), that is, corresponding to the  $\bar{q}_p \bar{q}_r \bar{q}_s$  polynomial products with  $prs = \{000\}, \{002\}, \{111\}, \{012\}$  (also with related harmonics at lower amplitude with  $n = 9, 10, 11$ ). The mode decomposition estimator (35) yields the quantitative constraint

$$F_{\text{NL}}^{\text{const}} = \frac{1}{N} \sum_n \bar{\alpha}_n^{\mathcal{R}\text{const}} \bar{\beta}_n^{\mathcal{R}\text{wmap}} = 35.1 \pm 27.4, \quad (f_{\text{NL}}^{\text{const}} = 149.4 \pm 116.8), \quad (40)$$

where  $F_{\text{NL}}$  is the bispectrum parameter normalised relative to the local model (16), while the lower case  $f_{\text{NL}}$  constraint employs the more model-dependent normalisation using the primordial shape function  $S(k, k, k) = 1$ . The variance here was determined from using 2000 Gaussian simulations in the same WMAP-realistic context. It is clear from this result that there is no evidence—given the present precision—for a significant constant primordial non-Gaussian signal.

### B. Local model

The canonical local shape, which we have already introduced in (39), covers a wide range of models where the non-Gaussianity is produced by local interactions. These models have their peak signal in “squeezed” states where one  $k_i$  is much smaller than the other two due to non-Gaussianity typically being produced on superhorizon scales. Single-field slow-roll inflation is dominated by the local shape, though  $f_{\text{NL}}^{\text{loc}}$  is tiny [20, 21]. The production of large non-Gaussianity during multiple field inflation [22–24] shows much greater promise of producing an observable signal through conversion of isocurvature into adiabatic perturbations. Large  $f_{\text{NL}}^{\text{loc}}$  can also be produced in curvaton models [25–27], at the end of inflation from reheating mechanisms [28] and also in more exotic scenarios such as (non-local)  $p$ -adic inflation [29] and the ekpyrotic scenario [30]. For more comprehensive references and recent examples please refer to the review, ref. [31].

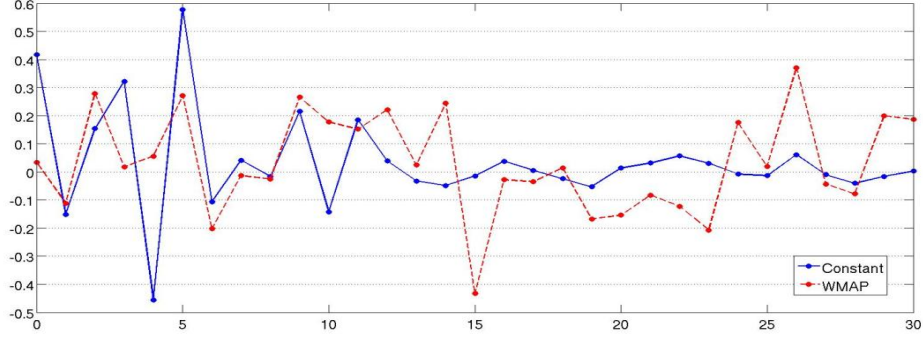


Figure 9: Comparison between constant model and recovered mode coefficients for the WMAP5 data. Note that the constant model incorporates features entirely due to the transfer functions (the acoustic peaks seen in modes  $n = 3, 4, 5$ ), which are indicators of its primordial origin.

The distinct mode decomposition of the local model has already been illustrated in fig. 3, together with the 3D CMB bispectrum in fig. 4 because they were used to validate this estimator methodology in section V. There are a number of differences from the constant model expansion reflecting the dominant signal along the edges of the tetrahedron, thus favouring higher order polynomials able to describe this localised signal. That is, as well as the periodic acoustic peak signal seen in the constant model ( $n = 3, 4, 5$ ), the spectrum is otherwise dominated by pure modes  $n = 9, 15, 26$  with  $prs = \{003\}, \{004\}, \{005\}$ . The expansion is not as rapidly convergent but the eigenmode partial sum achieves a 94% correlation by  $n = 31$ .

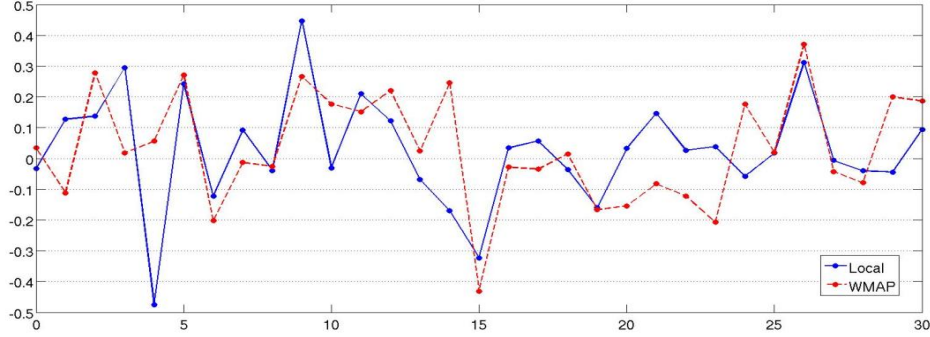


Figure 10: Comparison between local model expansion coefficients and recovered modes for the WMAP5 data. Note the relatively slow convergence of the local model and the apparent visual correlation of modes.

To aid comparison with the recovered WMAP bispectrum, we illustrate both in fig. 10. There appears to be some correlation between the two sets of data points which is reflected in the result from the mode estimator

$$F_{\text{NL}}^{\text{loc}} = 54.4 \pm 29.4 \quad (f_{\text{NL}}^{\text{loc}} = 54.4 \pm 29.4). \quad (41)$$

This nearly  $2\sigma$  result is consistent with that obtained by other groups in Table I. In particular, it can be compared to the raw result of  $f_{\text{NL}}^{\text{loc}} = 59 \pm 21$  obtained by the WMAP7 analysis, before marginalising over foregrounds. The similarities between the recovered WMAP bispectrum and the local bispectrum can be observed by comparing the 3D bispectrum in fig. 6 and fig. 4 respectively. There are obvious similarities around the edges of the tetrahedron where much of the local signal resides. However, we note an important precautionary point. Our analysis of the effects of the noise and mask before subtraction from simulations in a WMAP-realistic context indicates that these also have a very nearly ‘local’ shape. We believe the same is also true for likely foreground contaminants which we believe also contribute to the cubic term in the



estimator (12); this is indicated also by the significant effect of marginalisation over foregrounds on  $f_{\text{NL}}$  in the WMAP7 analysis. Our late-time analysis is obviously susceptible to all sources of non-Gaussianity, unlike the filtered primordial estimator searching for just a couple of modes. While such contaminants appear to have been largely removed from our analysis by the linear term (given their anisotropic nature and local shape), the local constraint (41) is clearly more susceptible to systematic effects. Further detailed investigations of these effects and the shape characterisation of noise, mask and contaminants using our mode expansions is the subject of a follow-up publication [15].

### C. Equilateral models

Bispectra dominated by contributions from nearly equilateral triangle configurations,  $k_1 \approx k_2 \approx k_3$  are produced through the amplification of nonlinear effects around the time modes exit the horizon, which can be achieved by modifying kinetic terms, as in the DBI model [32], or by explicitly adding higher derivative terms, such as in K-inflation [see, for example, 33]. For DBI inflation, this leads to non-Gaussianity being produced with a shape function of the form [32, 34]

$$S(k_1, k_2, k_3) = \frac{1}{k_1 k_2 k_3 (k_1 + k_2 + k_3)^2} \left[ \sum_i k_i^5 + \sum_{i \neq j} (2k_i^4 k_j - 3k_i^3 k_j^2) + \sum_{i \neq j \neq l} (k_i^3 k_j k_l - 4k_i^2 k_j^2 k_l) \right]. \quad (42)$$

This shape is illustrated in fig. 11, together with ghost inflation [35] and a third distinct single field equilateral shape found in a general analysis of such models [33]. Note that the generic equilateral shapes are not separable, but have been approximated to date using a separable ansatz commonly called the ‘equilateral model’ [11]:

$$S^{\text{equi}}(k_1, k_2, k_3) = \frac{1}{N} \frac{(k_2 + k_3 - k_1)(k_3 + k_1 - k_2)(k_1 + k_2 - k_3)}{k_1 k_2 k_3}. \quad (43)$$

Despite the apparent visual differences between these primordial shapes, particularly near the edges of the tetrahedral domain, the resulting CMB bispectra share at least a 95% or greater correlation ([see 5]). The CMB mode decomposition for these models is illustrated in fig. 12, showing very similar behaviour to the constant model also dominated by the acoustic peak coefficients  $n = 3, 4, 5$ . The resulting constraints from the modal estimator are:

$$\text{Equilateral: } F_{\text{NL}} = 25.1 \pm 26.4 \quad (f_{\text{NL}} = 143.5 \pm 151.2), \quad (44)$$

$$\text{DBI: } F_{\text{NL}} = 26.7 \pm 26.5 \quad (f_{\text{NL}} = 146.0 \pm 144.5), \quad (45)$$

$$\text{Ghost: } F_{\text{NL}} = 22.0 \pm 26.3 \quad (f_{\text{NL}} = 138.7 \pm 165.4), \quad (46)$$

$$\text{Single: } F_{\text{NL}} = 28.8 \pm 26.6 \quad (f_{\text{NL}} = 142.1 \pm 131.3). \quad (47)$$

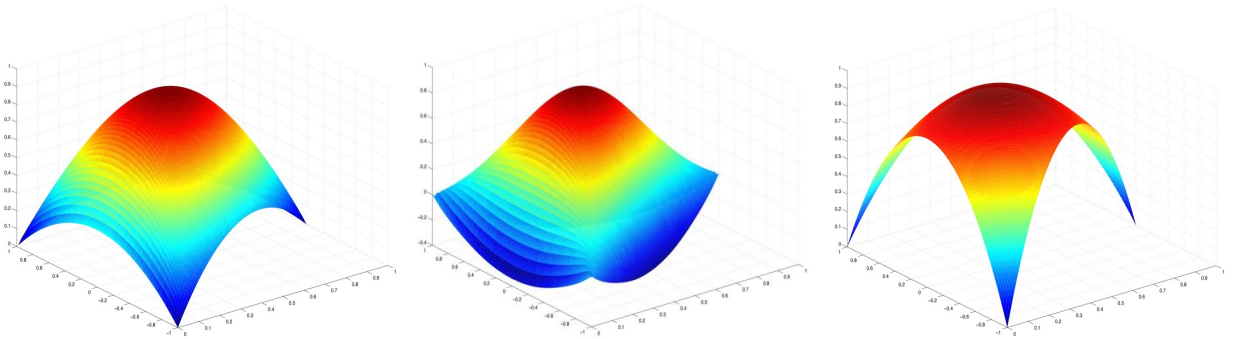


Figure 11: The shape function of models in the equilateral class which from left to right are DBI inflation, ghost inflation and the remaining single field inflation model.

Here, the local  $F_{\text{NL}}$  normalisation (16) yields much more consistent variances between models within the equilateral family than  $f_{\text{NL}}$  (as well as values comparable to local and other models). Note that there are up to 30% variations between the central values of these  $F_{\text{NL}}$  constraints despite the strong correlations between these bispectra; this is attributable to the different behaviour near the edges and faces where much of the apparent WMAP signal is localised. These results are consistent with the evolving constraints obtained in the literature to date, as shown in Table I.

Finally, we consider a separable ‘orthogonal’ shape  $S^{\text{orthog}}$  which is constructed from a linear combination of the constant and equilateral shape functions  $S^{\text{orthog}} \propto S^{\text{equil}} - 2/3$  (see [12, 36]). The constraint from the mode estimator (35) then becomes

$$F_{\text{NL}}^{\text{ortho}} = -16.3 \pm 27.3, \quad (f_{\text{NL}}^{\text{ortho}} = -79.4 \pm 133.3), \quad (48)$$

which is a less negative result than the latest WMAP7 limit  $f_{\text{NL}}^{\text{ortho}} = -199 \pm 104$ , but it remains consistent.

#### D. Flattened model

It is possible to consider inflationary vacuum states which are more general than the Bunch-Davies vacuum, such as an excited Gaussian (and Hadamard) state [37, see also discussions in Chen et al. 33, Meerburg et al. 36]. Observations of non-Gaussianity in this case might provide insight into trans-Planckian physics. The proposed non-separable shape for the bispectrum is

$$S^{\text{flat}}(k_1, k_2, k_3) \propto 6 \frac{k_1^2 + k_2^2 - k_3^2}{k_2 k_3} + 2 \text{ perms} + 2 \frac{k_1^2 + k_2^2 + k_3^2}{(k_1 + k_2 - k_3)^2 (k_2 + k_3 - k_1)^2 (k_3 + k_1 - k_2)^2}. \quad (49)$$

The bispectrum contribution from early times is dominated by flattened triangles, with e.g.  $k_3 \approx k_1 + k_2$ , and for a small sound speed  $c_s \ll 1$  can be large. Unfortunately, as the divergent analytic approximation

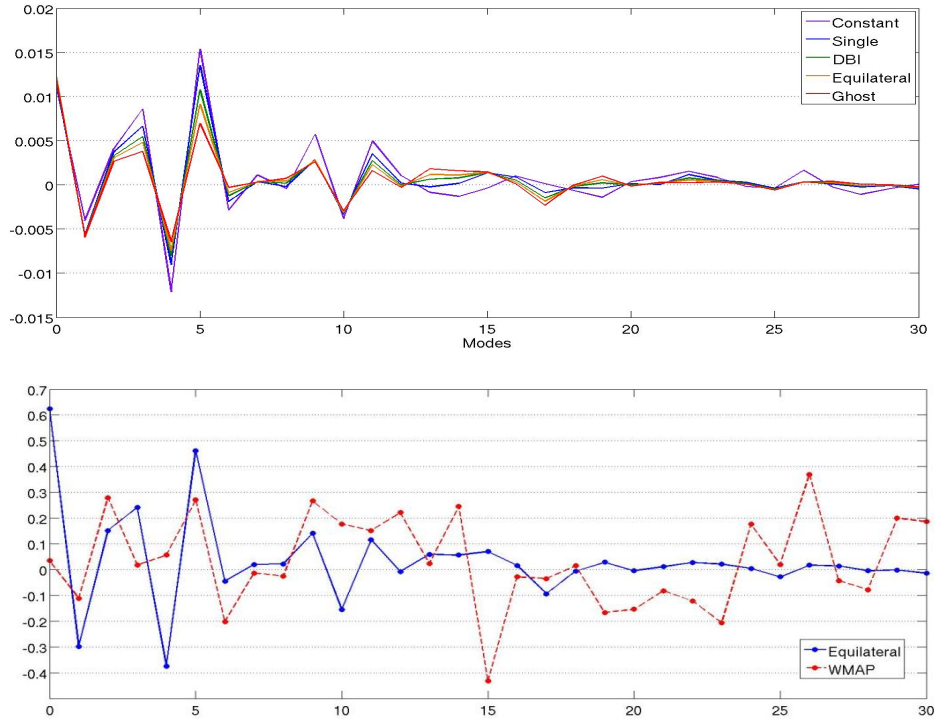


Figure 12: Equilateral model expansion coefficients  $\bar{\alpha}_n^{\mathcal{R}}$  compared between models (top panel) and compared with WMAP5 results (lower panel).

breaks down at the boundary of the allowed tetrahedron, some form of cut-off must be imposed, as shown for the smoothed shape in fig. 13 where an edge truncation has been imposed together with a mild Gaussian filter. This leads to a degree of predictive uncertainty, but the regularisation scheme ensures the primary signal is well-localised on the tetrahedral faces and is quite distinct from other separable shapes investigated to date (refer to ref. [5] for the specific details).

The resulting CMB spectrum reflects this behaviour with the dominant signal residing near the tetrahedral faces as shown in fig. 13. Figure 14 shows the flat model mode coefficients, which like the local model are only slowly convergent. Comparing the flat model with the coefficients obtained from WMAP, the mode estimator yields the new constraint:

$$F_{\text{NL}} = 35.4 \pm 29.2 \quad (f_{\text{NL}} = 18.1 \pm 14.9). \quad (50)$$

Despite the apparent visual similarity between the flat and WMAP CMB bispectra (figs 13 and 6) the present analysis does not reveal a significant correlation.

### E. Warm model

Finally, we consider warm inflation scenarios, that is, nearly scale-invariant models in which dissipative effects play a dynamical role, because these also may produce significant non-Gaussianity [38] (for a review see [39]). Contributions are again dominated by squeezed configurations but with a different more complex shape possessing a sign flip as the corner is approached. essentially making the signal orthogonal to the local shape. It can be shown that this makes the warm and local shapes essentially orthogonal with only a 33% correlation (see ref. [5] where the shape function and CMB bispectra are discussed). As with the flat model, uncertainties remain as to the validity of the approximations made as the corners and edges of the tetrapyd are approached. Comparison of the predicted warm bispectrum coefficients  $\bar{\beta}_n^{\text{warm}}$  with the WMAP data through the modal estimator (35) yields the constraint

$$F_{\text{NL}}^{\text{warm}} = 10.3 \pm 27.2 \quad (f_{\text{NL}}^{\text{warm}} = 47.4 \pm 125.4). \quad (51)$$

A previous WMAP3 warm inflation analysis obtained a lower central value  $f_{\text{NL}}^{\text{warm}} = -169 \pm 103$  [40] which is marginally consistent with (51) at the 95% confidence level. Probably the most significant difference is that the previous analysis did not include a linear term in the estimator (12) to account for noise and masking effects; these corrections are significant here as for the edge-dominated local model.

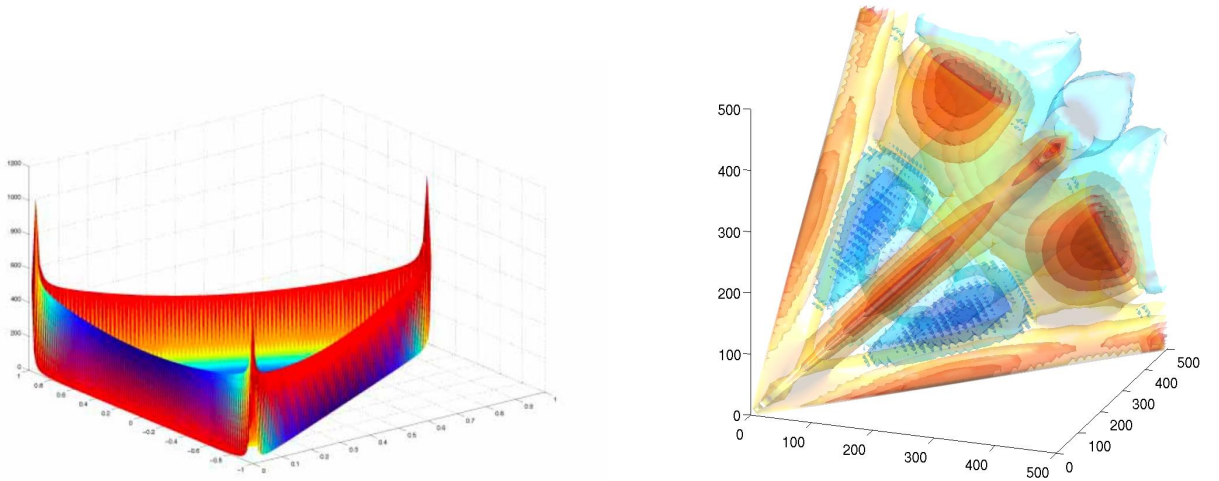


Figure 13: Flattened model: smoothed primordial shape function (left) and three-dimensional CMB bispectrum (right) for the flattened model.

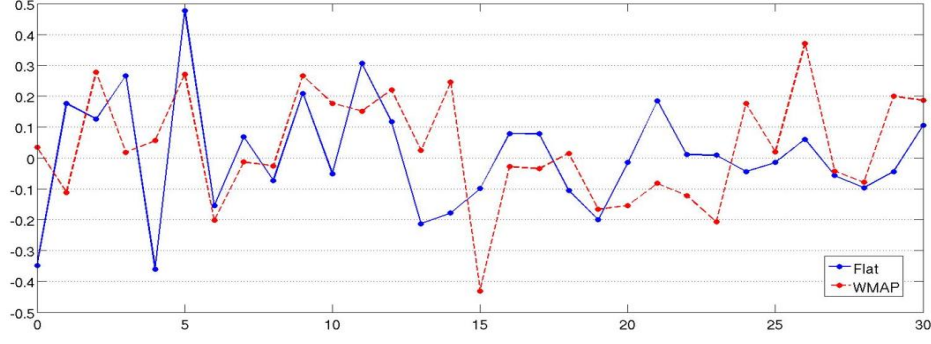


Figure 14: Flat model mode coefficients compared to WMAP5 mode coefficients.

### VIII. IMPLICATIONS FOR NON-SCALING FEATURE MODELS

It is possible to produce non-Gaussian signals which are not scale-invariant, such as models with a distinct feature in the inflaton potential. These usually take the form of either a step in the potential (models which have a long history, see e.g. ref. [41]) or those with a small oscillation superimposed onto the potential (which have become more popular recently, see e.g. ref. [42]). Two analytic forms for the resulting three point functions have been presented in ref. Chen et al. [41] with the expression we will analyse here taking

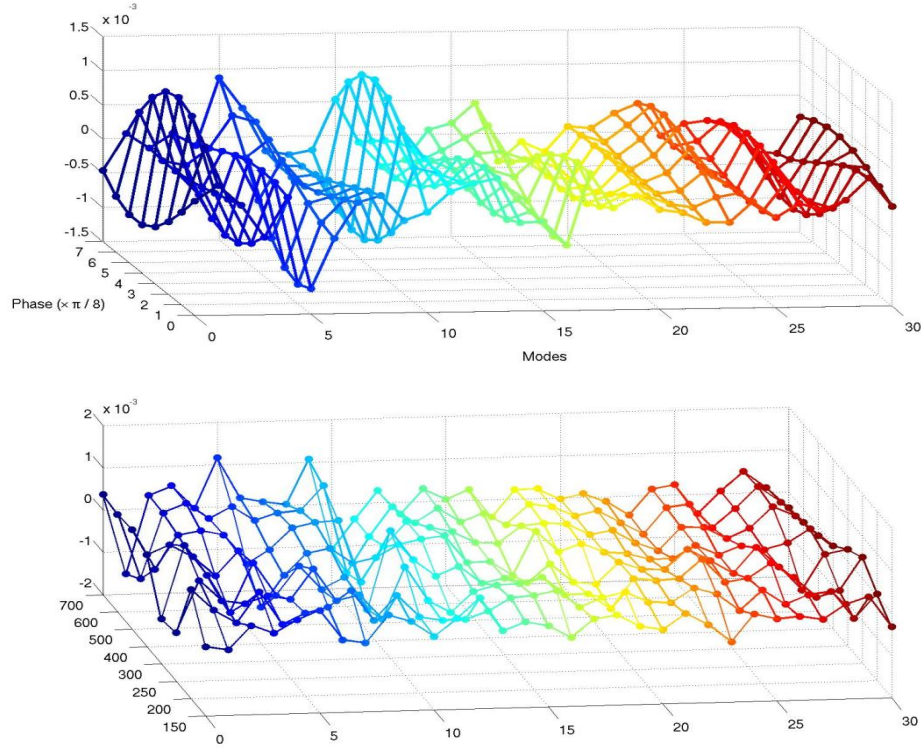


Figure 15: Feature model coefficients  $\bar{\alpha}_n^{\mathcal{R}}$  plotted in two-dimensions by mode number  $n$  and as function of phase  $\phi$  with  $l^* = 400$  (top panel) and as a function of scale  $l^*$  with  $\phi = 0$  (lower panel). Note how the characteristic  $n = 3, 4, 5$  primordial acoustic peak signature is affected (compare with fig. 9).

Scale Phase	150	200	250	300	400	500	600	700
0	57 (30)	-52 (33)	-25 (32)	1 (30)	1 (27)	8 (26)	18 (25)	23 (25)
$\pi/8$	67 (36)	-26 (27)	-36 (30)	-6 (25)	-4 (26)	-2 (27)	12 (26)	20 (25)
$\pi/4$	68 (42)	-10 (29)	-43 (30)	-11 (21)	-7 (25)	-10 (27)	-1 (28)	13 (27)
$3\pi/8$	49 (46)	7 (34)	-42 (32)	-18 (24)	-9 (25)	-14 (26)	-13 (28)	-2 (28)
$\pi/2$	15 (46)	32 (41)	-30 (35)	-32 (34)	-10 (25)	-16 (25)	-18 (27)	-14 (28)
$5\pi/8$	-19 (42)	63 (46)	-15 (35)	-38 (43)	-11 (25)	-16 (25)	-20 (26)	-20 (27)
$3\pi/4$	-39 (35)	87 (48)	0 (35)	-25 (41)	-11 (26)	-15 (25)	-21 (25)	-23 (26)
$7\pi/8$	-48 (30)	81 (43)	13 (34)	-11 (35)	-7 (27)	-13 (25)	-20 (25)	-23 (25)

Table II: Limits for a selection of feature models in the form  $F_{\text{NL}}$  (StDev).

the form

$$S^{\text{feat}}(k_1, k_2, k_3) = \frac{1}{N} \sin \left( 2\pi \frac{k_1 + k_2 + k_3}{3k^*} + \Phi \right), \quad (52)$$

where  $k^*$  is the associated with the physical scale of the feature in question and  $\Phi$  is an arbitrary phase factor. The alternative form with a logarithmic momentum dependence in the sin argument can be shown to be closely correlated with the simpler form (52), certainly on the present domain of study  $l_{\text{max}} = 500$ . Previously, we studied the shape and CMB bispectrum for a particular feature model (with  $k^* \approx l^*/\tau_0$  and  $l^* \approx 400$ ), showing that its non-scaling behaviour made it essentially independent of all the other shapes [5]. Such models can have starkly contrasting CMB bispectra as illustrated in fig. 18, disrupting the usual pattern of acoustic peaks which switch from correlation to anticorrelation on multipole scales  $l^*$ . Clearly, scale dependent feature models form a distinct category of bispectra beyond the equilateral, local, warm and flat families, so searches within WMAP and future data sets are well-motivated.

For the present WMAP5 analysis, we have studied the primordial feature shape (52) over a wide range of for which the CMB bispectra that we obtained could be accurately described by our  $n = 31$  eigenmodes, that is, for which we could obtain  $> 90\%$  convergence to  $b_{l_1 l_2 l_3}^{\text{feat}}$  for the partial sum (22). This restricted the scale parameters in (52) to the range  $l^* \geq 150$ , so we studied values  $l^* = 150, 200, 250, 300, 400, 500, 600, 700$ . For larger values  $l^* > 700$  the models became highly correlated with the constant model given that  $l_{\text{max}} = 500$ . No such restriction applied to the phase which was studied for each  $l^*$  over the full domain  $0 \leq \Phi < 2\pi$  in  $\pi/8$  steps (noting that models separated by  $\pi$  are merely anticorrelated). This entailed considerable computational effort calculating 64 distinct CMB bispectra at high accuracy using the robust methods previously described elsewhere [43]. The mode coefficients for the  $l^* = 400$  model are illustrated for the different phases in fig. 15, demonstrating how the characteristic acoustic peak signal in  $n = 3, 4, 5$  can be modified (compare the constant model fig. 9). The strong dependence of the mode coefficients on the different multipole scales  $l^*$  (at fixed phase  $\Phi = 0$ ) are shown in fig. 15.

Results from the modal estimator for all the feature models investigated are provided in Table II. Note that the constraints are given in terms of the normalised quantity  $F_{\text{NL}}$  defined in (16), since there is no simple generalisation of the primordial normalisation used for  $f_{\text{NL}}$  without scale-invariance. As before, the variances (given in parentheses) are those obtained for the same set of models from 1000 Gaussian simulations. The results are illustrated graphically in fig. 16 showing the relative significance of the central  $F_{\text{NL}}$  values relative to the standard deviation. The result with the highest significance is that for the feature model with  $l^* = 150$  and zero phase which achieves a  $1.9\sigma$  significance. The 3D bispectrum for this model is shown in fig. 18 demonstrating how such models can reproduce the apparent scale-dependence observed in the WMAP bispectrum (see fig. 6). However, we note that this model is close to the resolution limit set by the eigenmodes deployed (like the other cases of higher significance). The results over the full domain of feature models investigated remain consistent with the Gaussian hypothesis with no significant detection found on the WMAP domain for  $l \leq 500$ .



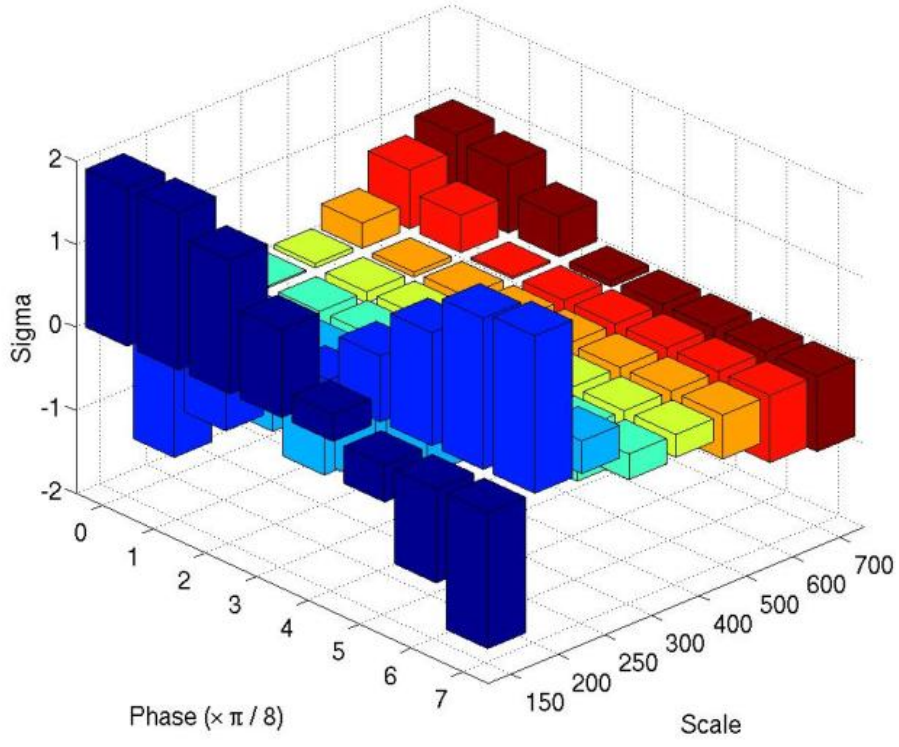


Figure 16: Significance of feature model bispectra  $F_{\text{NL}}/\Delta F_{\text{NL}}$  using WMAP data with the modal estimator (35). This is plotted as a function of the multipole scale  $l^*$  and the phase of feature models given by (52).

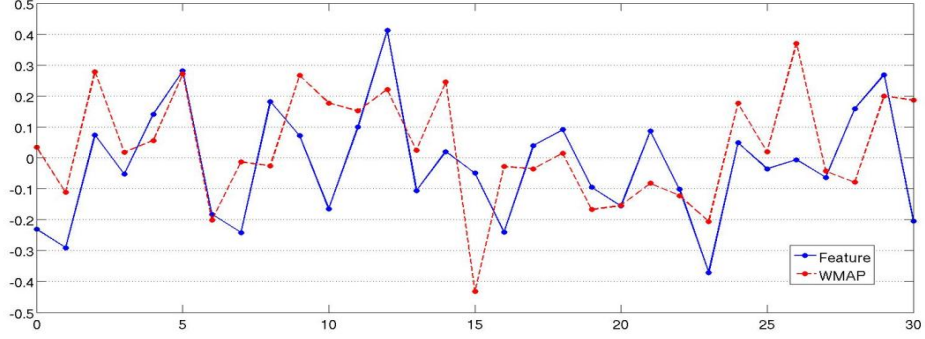


Figure 17: Best fit feature model coefficients ( $l^* = 150$ ,  $\phi = 0$ ) compared to WMAP5 mode coefficients.

## IX. TOWARDS A MEASURE OF THE TOTAL INTEGRATED BISPECTRUM $\bar{F}_{\text{NL}}$

Our focus in this paper has been on recovering the observed bispectrum  $b_{l_1 l_2 l_3}$  which contains more information than  $f_{\text{NL}}^{\text{th}}$  constraints for particular models. We can also consider squaring this quantity and

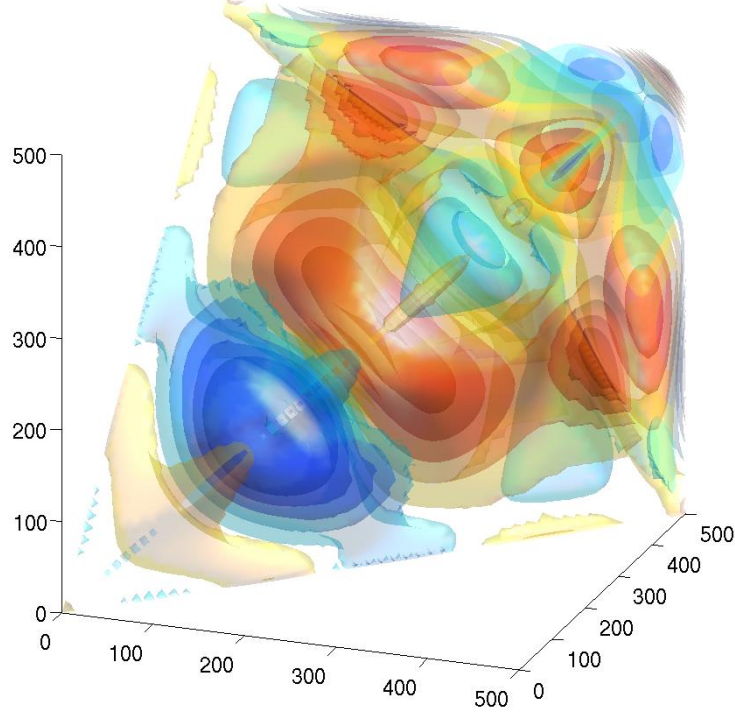


Figure 18: Three-dimensional CMB bispectrum calculated for the best fit feature model ( $l^* = 150$ ,  $\phi = 0$ ) . Note how the scale-dependence of the central peaks mimics at some level that observed in the WMAP data.

summing over all multipoles to obtain a total integrated nonlinearity parameter  $\bar{F}_{\text{NL}}$  defined by

$$\bar{F}_{\text{NL}}^2 \equiv \frac{1}{N_{\text{loc}}^2} \sum_{l_i m_i} \frac{B_{m_1 m_2 m_3}^{l_1 l_2 l_3}}{C_{l_1} C_{l_2} C_{l_3}}^2 = \frac{1}{N_{\text{loc}}^2} \sum_{l_i} \frac{h_{l_1 l_2 l_3}^2 b_{l_1 l_2 l_3}^2}{C_{l_1} C_{l_2} C_{l_3}}. \quad (53)$$

Substituting our mode decomposition (30) into the expression for the integrated bispectrum we can find the leading order contribution from the three-point correlator to be<sup>1</sup>

$$\begin{aligned} \bar{F}_{\text{NL}}^2|_{3\text{pt}} &\equiv \frac{1}{N_{\text{loc}}^2} \sum_{l_i} \frac{h_{l_1 l_2 l_3}^2 b_{l_1 l_2 l_3}^2}{C_{l_1} C_{l_2} C_{l_3}} = \frac{1}{N_{\text{loc}}^2} \sum_{l_i} w_s(l_1, l_2, l_3) \left( \sum_n \bar{\beta}_n^{\mathcal{R}} \bar{\mathcal{R}}_n \right) \left( \sum_p \bar{\beta}_p^{\mathcal{R}} \bar{\mathcal{R}}_p \right) \\ &\approx \frac{1}{N_{\text{loc}}^2} \sum_n \sum_p \bar{\beta}_n^{\mathcal{R}} \bar{\beta}_p^{\mathcal{R}} \langle \bar{\mathcal{R}}_n, \bar{\mathcal{R}}_p \rangle \approx \frac{\sum_n \bar{\beta}_n^{\mathcal{R}2}}{\sum_n \bar{\alpha}_n^{\mathcal{R}2}}. \end{aligned} \quad (55)$$

Unfortunately, however, the expectation value  $\langle \bar{F}_{\text{NL}}^2 \rangle$  contains much more than the just contributions from the three-point correlator. There is necessarily a contribution from products of two-point correlators,

<sup>1</sup> As an unambiguous signature of a significant bispectrum we should compare  $\bar{F}_{\text{NL}}$  with the skewness  $\gamma_1$  which is given by [44]

$$\gamma_1 \equiv \left\langle \left( \frac{\Delta T}{T}(\hat{\mathbf{n}}) \right)^3 \right\rangle = \frac{1}{4\pi} \sum_{l_i} h_{l_1 l_2 l_3}^2 b_{l_1 l_2 l_3}. \quad (54)$$

In principle, the skewness can conspire to vanish even with a non-zero bispectrum  $b_{l_1 l_2 l_3}$  because it is not positive definite, in contrast to the bispectrum contribution to  $\bar{F}_{\text{NL}}$ .

reflecting both cosmic variance and instrument noise and there are also higher-order contributions, which are derived in the Appendix. The leading order terms are found to be

$$\bar{F}_{\text{NL}}^2 \approx \frac{1}{N_{\text{loc}}^2} \left( 6n_{\text{max}} + \sum_n^{n_{\text{max}}} [F_{\text{NL}}^2 \bar{\alpha}_n^{\kappa 2} + \langle \bar{\beta}_n^{\kappa 2} \rangle_6 + \dots] \right), \quad (56)$$

where the first term represents the underlying variance, the second is the integrated bispectrum we seek, and the third term arises from contributions from the four-point correlator or trispectrum  $T_{l_3 l_4}^{l_1 l_2}(L)$ .

The non-Gaussian corrections from the bispectrum and trispectrum (and possibly higher correlators) to  $\bar{F}_{\text{NL}}^2$  are not easily distinguishable. Nevertheless, we know the Gaussian expectation so it can still be used to measure any deviation from Gaussianity, by determining the amplitude and nature of any differences. To prove the utility of this statistic we have applied it to both Gaussian and local simulations. As we have shown in a previous paper [44] our local map simulations contain a negligible trispectrum and so we can use the approximation

$$\bar{F}_{\text{NL}}^2 \approx \langle \bar{F}_{\text{NL}}^{G^2} \rangle + \frac{1}{N_{\text{loc}}^2} F_{\text{NL}}^2 \sum_n \bar{\alpha}_n^{\kappa 2}, \quad (57)$$

where we have defined  $\langle \bar{F}_{\text{NL}}^{G^2} \rangle$  to be the value recovered from Gaussian simulations. As  $N^2 = \sum_n^{n_{\text{max}}} \bar{\alpha}_n^{\kappa 2}$  and we normalise all models to have  $N = N^{\text{local}}$  we can then attempt to recover  $\bar{F}_{\text{NL}}$  without assuming any particular form for  $\bar{\alpha}_n^{\kappa}$ . The recovered  $\bar{F}_{\text{NL}}$  will then be given by the following formula

$$F_{\text{NL}}^{\text{rec}} = \sqrt{\bar{F}_{\text{NL}}^2 - \langle \bar{F}_{\text{NL}}^{G^2} \rangle}. \quad (58)$$

We estimated  $\langle \bar{F}_{\text{NL}}^{G^2} \rangle$  from 1000 Gaussian simulations and then calculated  $F_{\text{NL}}^{\text{rec}}$  from 100 simulations with various input  $F_{\text{NL}}$ . The results are presented in table III and are plotted as cumulative sum of squared mode coefficients in figure 19. The encouraging results show that the statistic is recovering the input  $f_{\text{NL}}$  to a reasonable degree of accuracy, although with a slight tendency to under-estimation. The biggest surprise is that the error bars for this general bispectrum estimator only increase by a factor of 50% from those when the particular local form is assumed explicitly. We note, however, that (57) gives rise to a  $\chi^2$ -distribution, so we have to take care in assuming Gaussianity for small  $n_{\text{max}}$ . Again, we will further explore the utility of such general modal statistics elsewhere [15].

Input $f_{\text{NL}}$	Mean	StDev	Null
50	55.86	48.53	36
75	65.39	46.95	20
100	95.28	41.61	8
150	138.8	44.87	3
200	187.39	41.55	0

Table III:  $F_{\text{NL}}^{\text{rec}}$  as recovered from 100 simulated local maps. Null refers to the number of maps in which the recovered  $F_{\text{NL}}$  is less than the Gaussian expectation value.

With the efficacy of the  $\bar{F}_{\text{NL}}$  statistic established we have also applied it to the WMAP5 data. This yields the unexpected result that  $F_{\text{NL}}$  obtained from the WMAP5 data is less than that which we would expect from a typical Gaussian map by slightly over  $1\sigma$  (see the cumulative sum in fig. 20). This is somewhat surprising because one would expect a late-time estimator to be susceptible to foregrounds or other contamination, but the deviation remains statistically insignificant. In principle, this result could be due to a large negative trispectrum or higher order contribution (see Appendix). However, neglecting this possibility, the result shown in fig. 20 indicates that there is no significant contribution to the bispectrum from the first 31 eigenmodes. This would constrain virtually all smooth scale invariant shapes, as well as the feature models we have surveyed. The only remaining possibility for a bispectrum detection (at the present precision) would then be for oscillatory models with sufficiently high frequencies or bispectra with particularly sharp, or localised, features (i.e. those which require  $n > 31$  for an accurate description). We have good evidence, therefore, for the null hypothesis that we live in a Gaussian universe.



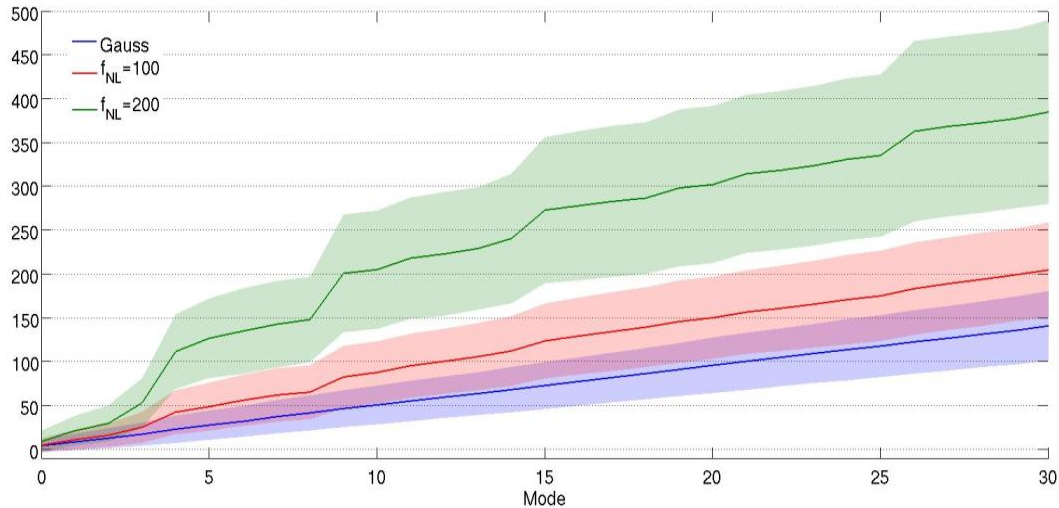


Figure 19: Cumulative sum of mode contributions to the total  $\bar{F}_{\text{NL}}^2$  (55) for the local  $F_{\text{NL}} = 100$  (red) and  $F_{\text{NL}} = 200$  (green) map simulations compared with Gaussian maps (blue). The  $1\sigma$  variance is shaded around the mean value obtained from 100 simulations (1000 simulations for the Gaussian case).

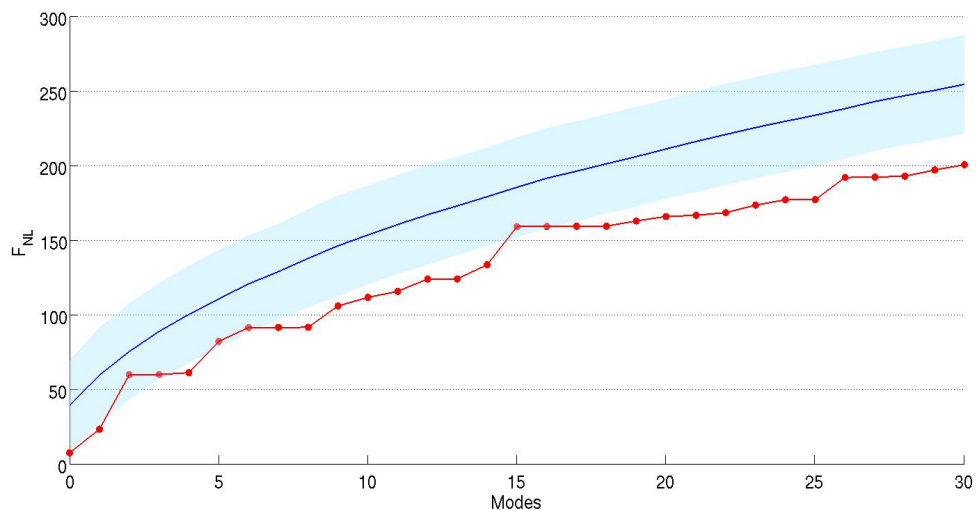


Figure 20: Cumulative sum of mode contributions to the total  $\bar{F}_{\text{NL}}$  (55) for the WMAP data compared with Gaussian map simulations as in fig. 19.

## X. DISCUSSION AND CONCLUSIONS

We have implemented and validated separable mode expansions with a general late-time CMB bispectrum estimator, using it to investigate a wide range of primordial models with WMAP 5-year data. Notable new constraints include those on non-scaling feature models, trans-Planckian (flat) models and warm inflation. The results for nearly scale-invariant models are summarised in Table IV, demonstrating consistency with previous constraints on equilateral and local models. Note that we adopt a nonlinearity parameter  $F_{\text{NL}}$  normalised to facilitate direct comparison between the local  $f_{\text{NL}}$  and any other model. We found no evidence for significant deviations from Gaussianity for any specific model (at 95% confidence). Feature models were surveyed over a wide range of parameters with periodicities above  $l^* = 150$  and over the full domain of phase values. Again, no significant bispectrum detection was made, though given the nature of this survey

Model	$F_{\text{NL}}$	$(f_{\text{NL}})$
<b>Constant</b>	$35.1 \pm 27.4$	$(149.4 \pm 116.8)$
<b>DBI</b>	$26.7 \pm 26.5$	$(146.0 \pm 144.5)$
<b>Equilateral</b>	$25.1 \pm 26.4$	$(143.5 \pm 151.2)$
<b>Flat (Smoothed)</b>	$35.4 \pm 29.2$	$(18.1 \pm 14.9)$
<b>Ghost</b>	$22.0 \pm 26.3$	$(138.7 \pm 165.4)$
<b>Local</b>	$54.4 \pm 29.4$	$(54.4 \pm 29.4)$
<b>Orthogonal</b>	$-16.3 \pm 27.3$	$(-79.4 \pm 133.3)$
<b>Single</b>	$28.8 \pm 26.6$	$(142.1 \pm 131.3)$
<b>Warm</b>	$24.2 \pm 27.3$	$(94.7 \pm 106.8)$

Table IV: Limits for all known scale invariant models

some models provide a better a posteriori fit to the data than others.

As we have emphasised throughout, more information can be extracted from the mode decomposition of the data than a few  $F_{\text{NL}}$ 's for specific models. Given that we have constructed a complete orthonormal basis  $\bar{\mathcal{R}}_n$  we can use the mode coefficients  $\bar{\beta}_n^{\mathcal{R}}$  to directly reconstruct the full CMB bispectrum using the partial sum (30). We plotted the result for WMAP5 in fig. 6 which, despite its low significance, revealed interesting qualitative features similar to the local model (4), but without the periodicity expected from acoustic peaks. We discussed a positive-definite measure for the total integrated bispectrum constructed from the mode coefficients  $\bar{F}_{\text{NL}}^2 = \sum_n \bar{\beta}_n^{\mathcal{R}2}$ , which was used to recover  $f_{\text{NL}}$  from map simulations in a model independent manner (though with larger variance). For WMAP5 data the integrated  $\bar{F}_{\text{NL}}$  was found to be small and again consistent with a Gaussian hypothesis.

Despite the absence of any convincing evidence for a statistically significant CMB bispectrum in the present analysis, many avenues remain open for further investigation using the present methodology. The late-time modal estimator (35) can identify any bispectrum whether generated at early times like inflation or sourced since decoupling by cosmic strings, gravitational lensing, or second-order gravitational effects. Unlike the primordial estimator, the general mode expansion can also be used to characterise noise and foregrounds, which need to be identified and subtracted through the linear term in the estimator (12). The efficacy of this removal and other validation checks which may affect a residual local signal will be published shortly [15]. Finally, we note again that these methods can be pressed much further with existing and future data, especially from Planck. The anticipated Planck variance  $\Delta f_{\text{NL}} \approx 5$  will substantially improve sensitivity to specific bispectrum shapes, leaving significant discovery potential available in the near future. We note also that these separable mode techniques have been adapted for general CMB trispectrum estimation, in principle, making tractable the investigation of all planar primordial trispectra [44]. Analogous methods can also be applied to modal bispectrum extraction for large-scale structure and in other contexts. For the time being, however, this general bispectrum survey uncovers no significant evidence of non-Gaussianity which would undermine the standard predictions of the simplest models of inflation.

## XI. ACKNOWLEDGEMENTS

We are very grateful for many informative discussions with Donough Regan, Xingang Chen, Anthony Challinor and Alessandro Renzi. Simulations were performed on the COSMOS supercomputer (an Altix 4700) which is funded by STFC, HEFCE and SGI. We are particularly grateful for computer support from Andrey Kaliazin. JRF, ML and EPS were supported by STFC grant ST/F002998/1 and the Centre for Theoretical Cosmology. EPS is grateful for the hospitality of the Arnold Sommerfeld Centre and the Universe Excellence Cluster in Munich.

## XII. APPENDIX - HIGHER-ORDER CONTRIBUTIONS TO $F_{\text{NL}}^2$

Consider the expectation value of the quantity  $\langle \bar{F}_{\text{NL}}^2 \rangle = \sum_n \langle \bar{\beta}_n^{\mathcal{R}2} \rangle$  which we defined in (53) and discussed in section IX. Our definition of the bispectrum mode coefficients  $\bar{\beta}_n^{\mathcal{R}}$  in (34) can be expanded using the expressions in (31) to the explicit form

$$\bar{\beta}_n^{\mathcal{R}} = \sum_{l_i} \frac{h_{l_1 l_2 l_3} \bar{\mathcal{R}}_n(l_1, l_2, l_3)}{v_{l_1} v_{l_2} v_{l_3} \sqrt{C_{l_1} C_{l_2} C_{l_3}}} \sum_{m_i} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3}. \quad (59)$$

It is instructive at this point to repeat the derivation of the expectation value of the coefficient  $\langle \bar{\beta}_n^{\mathcal{R}} \rangle$  for an ensemble of universes with both a given  $F_{\text{NL}}$  as in (16) and a given theoretical bispectrum  $b_{l_1 l_2 l_3}^{\text{th}(F_{\text{NL}}=1)}$  (for further details see ref. [1]). We describe the theoretical bispectrum  $b_{l_1 l_2 l_3}^{\text{th}(F_{\text{NL}}=1)}$  by the orthonormal mode expansion coefficients  $\bar{\alpha}_n^{\mathcal{R}}$  as in (30) or, equivalently, the separable modes  $\bar{\alpha}_n^{\mathcal{S}}$  as in (22), distinguishing it from the observed bispectrum  $b_{l_1 l_2 l_3}$  described by  $\bar{\beta}_n^{\mathcal{R}}$  or  $\bar{\beta}_n^{\mathcal{S}}$ . Substituting the separable mode expansion of the bispectrum (22) into the expression for the recovered coefficient (34) we find

$$\langle \bar{\beta}_n^{\mathcal{S}} \rangle = \int d^2 \hat{\mathbf{n}} \langle \bar{M}_{\{p\}}(\hat{\mathbf{n}}) \bar{M}_r(\hat{\mathbf{n}}) \bar{M}_{\{s\}}(\hat{\mathbf{n}}) \rangle = \sum_{l_i, m_i} \frac{\bar{q}_{\{p\}}(l_1) \bar{q}_r(l_2) \bar{q}_{\{s\}}(l_3)}{v_{l_1} v_{l_2} v_{l_3} \sqrt{C_{l_1} C_{l_2} C_{l_3}}} \left( \mathcal{G}_{m_1 m_2 m_3}^{l_1 l_2 l_3} \right)^2 \langle b_{l_1 l_2 l_3} \rangle \quad (60)$$

$$= \sum_{l_i, m_i} \frac{h_{l_1 l_2 l_3}^2 \bar{\mathcal{Q}}_n(l_1, l_2, l_3)}{v_{l_1} v_{l_2} v_{l_3} \sqrt{C_{l_1} C_{l_2} C_{l_3}}} \sum_p F_{\text{NL}} \bar{\alpha}_p^{\mathcal{S}} \bar{\mathcal{Q}}_p(l_1, l_2, l_3) \frac{\sqrt{C_{l_1} C_{l_2} C_{l_3}}}{v_{l_1} v_{l_2} v_{l_3}} \quad (61)$$

$$= F_{\text{NL}} \sum_p \bar{\alpha}_p^{\mathcal{S}} \sum_{l_1 l_2 l_3} w_s \bar{\mathcal{Q}}_n \bar{\mathcal{Q}}_p = F_{\text{NL}} \sum_p \bar{\gamma}_{np} \bar{\alpha}_p^{\mathcal{S}}. \quad (62)$$

From (29) we can transform this into the orthonormal basis  $\bar{\mathcal{R}}_n$  to find the simple result

$$\langle \bar{\beta}_n^{\mathcal{R}} \rangle = F_{\text{NL}} \bar{\alpha}_n^{\mathcal{R}}. \quad (63)$$

Here we have ignored the linear term in (34) because its expectation value vanishes.

Now the expectation value of the square of this coefficient  $\bar{\beta}_n^{\mathcal{R}}$  necessarily involves the six-point function, because the expression takes the form

$$\langle \bar{\beta}_n^{\mathcal{R}2} \rangle = \sum \frac{h_{l_1 l_2 l_3} h_{l_4 l_5 l_6} \bar{\mathcal{R}}_n(l_1, l_2, l_3) \bar{\mathcal{R}}_n(l_4, l_5, l_6)}{v_{l_1} v_{l_2} v_{l_3} v_{l_4} v_{l_5} v_{l_6} \sqrt{C_{l_1} C_{l_2} C_{l_3} C_{l_4} C_{l_5} C_{l_6}}} \times \sum_{m_i} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_4 & l_5 & l_6 \\ m_4 & m_5 & m_6 \end{pmatrix} \langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} a_{l_4 m_4} a_{l_5 m_5} a_{l_6 m_6} \rangle. \quad (64)$$

Here we note that we can include a cut-sky, as well as noise and beam effects, as we did previously in (14) and (13) respectively.

The expectation value of the six-point function has a variety of non-vanishing contributions from combi-

nations of lower order correlators, which become (after summing over equivalent permutations):

$$\begin{aligned}
\langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} a_{l_4 m_4} a_{l_5 m_5} a_{l_6 m_6} \rangle = & 6 \delta_{l_1 l_4} \delta_{l_2 l_5} \delta_{l_3 l_6} \delta_{m_1 - m_4} \delta_{m_2 - m_5} \delta_{m_3 - m_6} C_{l_1} C_{l_2} C_{l_3} \\
& + 9 \delta_{l_1 l_2} \delta_{l_3 l_4} \delta_{l_5 l_6} \delta_{m_1 - m_2} \delta_{m_3 - m_4} \delta_{m_5 - m_6} C_{l_1} C_{l_3} C_{l_5} \\
& + \mathcal{G}_{m_1 m_2 m_3}^{l_1 l_2 l_3} \mathcal{G}_{m_4 m_5 m_6}^{l_4 l_5 l_6} b_{l_1 l_2 l_3} b_{l_4 l_5 l_6} \\
& + 9 \mathcal{G}_{m_1 m_2 m_4}^{l_1 l_2 l_4} \mathcal{G}_{m_3 m_5 m_6}^{l_3 l_5 l_6} b_{l_1 l_2 l_4} b_{l_3 l_5 l_6} \\
& + 6 \delta_{l_1 l_2} \delta_{m_1 - m_2} C_{l_1} \sum_{LM} (-1)^M \begin{pmatrix} l_3 & l_4 & L \\ m_3 & m_4 & M \end{pmatrix} \begin{pmatrix} l_5 & l_6 & L \\ m_5 & m_6 & -M \end{pmatrix} T_{l_5 l_6}^{l_3 l_4}(L) \\
& + 9 \delta_{l_1 l_4} \delta_{m_1 - m_4} C_{l_1} \sum_{LM} (-1)^M \begin{pmatrix} l_2 & l_3 & L \\ m_2 & m_3 & M \end{pmatrix} \begin{pmatrix} l_5 & l_6 & L \\ m_5 & m_6 & -M \end{pmatrix} T_{l_5 l_6}^{l_2 l_3}(L) \\
& + \text{higher order terms}, \tag{65}
\end{aligned}$$

where the CMB trispectrum  $T_{l_3 l_4}^{l_1 l_2}(L)$  is reviewed and discussed at some length in a recent companion paper [44].

Consecutively labelling the  $i$ th term in the six-point expression (65) above, we now evaluate each specific contribution to the expectation value of  $\bar{F}_{\text{NL}}^2$ , denoting modes term-by-term as  $\langle \bar{\beta}_n^{\mathcal{R}^2} \rangle_i$ . Recall that we have orthonormal basis eigenmodes  $\bar{\mathcal{R}}_n$  (26) which satisfy

$$\sum_{l_i} \frac{h_{l_1 l_2 l_3}^2 \bar{\mathcal{R}}_n(l_1, l_2, l_3) \bar{\mathcal{R}}_p(l_1, l_2, l_3)}{v_{l_1}^2 v_{l_2}^2 v_{l_3}^2} = \sum_{l_i} w_s(l_1, l_2, l_3) \bar{\mathcal{R}}_n(l_1, l_2, l_3) \bar{\mathcal{R}}_p(l_1, l_2, l_3) = \langle \bar{\mathcal{R}}_n, \bar{\mathcal{R}}_p \rangle = \delta_{np}. \tag{66}$$

The first term from (65) is a product of two-point correlators  $C_l$  in the numerator which cancels with the weighting in the denominator of (64) to become simply

$$\langle \bar{\beta}_n^{\mathcal{R}^2} \rangle_1 = 6 \langle \bar{\mathcal{R}}_n, \bar{\mathcal{R}}_n \rangle = 6. \tag{67}$$

This is the primary Gaussian noise contribution which cumulatively dominates  $\bar{F}_{\text{NL}}^2$ , as discussed in section IX, and which has been confirmed quantitatively in Gaussian simulations. The second term from (65) also consists of products of two-point correlators which divide out, but it does not simply further:

$$\langle \bar{\beta}_n^{\mathcal{R}^2} \rangle_2 = 9 \sum_{l_i} \frac{h_{l_1 l_1 l_3} h_{l_3 l_5 l_5}}{v_{l_1}^2 v_{l_3}^2 v_{l_5}^2} \bar{\mathcal{R}}_n(l_1, l_1, l_3) \bar{\mathcal{R}}_n(l_3, l_5, l_5) \sum_{m_i} \begin{pmatrix} l_1 & l_1 & l_3 \\ m_1 & -m_1 & 0 \end{pmatrix} \begin{pmatrix} l_3 & l_5 & l_5 \\ 0 & m_5 & -m_5 \end{pmatrix}. \tag{68}$$

Despite the difficulty evaluating this expression explicitly, it appears that the product of Wigner-3j symbols will behave asymptotically as  $l^{-1}$ , so we can expect the summed product of the  $\bar{\mathcal{R}}_n$  eigenfunctions in (68) to be significantly suppressed relative to the inner product in (67). For  $l_{\text{max}} \gg 1$ , therefore, we expect  $\langle \bar{\beta}_n^{\mathcal{R}^2} \rangle_1 \gg \langle \bar{\beta}_n^{\mathcal{R}^2} \rangle_2$ . For small non-Gaussianity, it may be necessary to calculate this term explicitly, though it is more costly to evaluate than the usual inner product (21). Alternatively, its effect can be determined from Gaussian simulations, which already confirm the clear dominance of the first contribution (67).

The third contribution from (65) is the straightforward product of the bispectra sought in IX, which from (30) simply collapses to

$$\begin{aligned}
\langle \bar{\beta}_n^{\mathcal{R}^2} \rangle_3 &= \left[ \sum_{l_i} \frac{h_{l_1 l_2 l_3}^2}{v_{l_1}^2 v_{l_2}^2 v_{l_3}^2 \sqrt{C_{l_1} C_{l_2} C_{l_3}}} \bar{\mathcal{R}}_n(l_1, l_2, l_3) b_{l_1 l_2 l_3} \right]^2 = \left[ \sum_{l_i} w_s(l_1, l_2, l_3) \bar{\mathcal{R}}_n F_{\text{NL}} \sum_p \bar{\alpha}_n^{\mathcal{R}} \bar{\mathcal{R}}_n \right]^2 \\
&= F_{\text{NL}}^2 \bar{\alpha}_n^{\mathcal{R}^2}, \tag{69}
\end{aligned}$$

where here we have distinguished the recovered  $b_{l_1 l_2 l_3}$  in the above from that normalised with  $F_{\text{NL}} = 1$  which defines the  $\bar{\alpha}_n^{\mathcal{R}}$  coefficients in (30). The fourth contribution also arises from products of the bispectrum but in the form

$$\langle \bar{\beta}_n^{\mathcal{R}^2} \rangle_4 = 9 F_{\text{NL}}^2 \sum_L \frac{1}{2L+1} \left[ \sum_p \bar{\alpha}_p^{\mathcal{R}} \sum_{l_1 l_2} w_s(l_1, l_2, L) \bar{\mathcal{R}}_n((l_1, l_2, L)) \bar{\mathcal{R}}_p((l_1, l_2, L)) \right]^2. \tag{70}$$

However, as for the second term (68), it is clear that the additional weighting  $(2L+1)^{-1}$  will generically suppress this contribution, so that  $\langle \bar{\beta}_n^{\kappa^2} \rangle_4 \ll \langle \bar{\beta}_n^{\kappa^2} \rangle_3$  for  $l_{\max} \gg 1$ . Finally, we consider terms involving combinations of the two-point and four-point correlators which simplify considerably because of the following identities for summed Wigner-3j symbols:

$$\begin{aligned} \sum_{m_2, m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_2 & l_3 & l_4 \\ m_2 & m_3 & m_4 \end{pmatrix} &= \frac{\delta_{l_1 l_4} \delta_{m_1 - m_4}}{2l_1 + 1}, \\ \sum_{m_1} (-1)^{-l_1 - m_1} \begin{pmatrix} l_1 & l_1 & l_3 \\ m_1 & -m_1 & 0 \end{pmatrix} &= \sqrt{2l_1 + 1} \delta_{l_3 0}. \end{aligned} \quad (71)$$

The fifth term from (65) reduces to

$$\langle \bar{\beta}_n^{\kappa^2} \rangle_5 = 6 \sum_{l_i} \frac{\sqrt{(2l_1+1)(2l_2+1)(2l_3+1)(2l_4+1)}}{4\pi v_{l_1}^2 v_{l_2} v_{l_3} v_{l_4} \sqrt{C_0 C_{l_2} C_{l_3} C_{l_4}}} \begin{pmatrix} l_2 & l_3 & l_4 \\ 0 & 0 & 0 \end{pmatrix} \bar{\mathcal{R}}_n(l_1, l_1, 0) \bar{\mathcal{R}}_n(l_2, l_3, l_4) T_{l_3 l_4}^{0 l_2}(l_2), \quad (72)$$

which is a monopole term which can be ignored. The sixth term from (65) can be expressed in the form

$$\begin{aligned} \langle \bar{\beta}_n^{\kappa^2} \rangle_6 &= 9 \sum_L \frac{1}{v_L^2} \sum_{l_i} \frac{\sqrt{(2l_1+1)(2l_2+1)(2l_3+1)(2l_4+1)}}{4\pi v_{l_1} v_{l_2} v_{l_3} v_{l_4} \sqrt{C_{l_1} C_{l_2} C_{l_3} C_{l_4}}} \begin{pmatrix} L & l_1 & l_2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} L & l_3 & l_4 \\ 0 & 0 & 0 \end{pmatrix} \\ &\quad \times \bar{\mathcal{R}}_n(L, l_1, l_2) \bar{\mathcal{R}}_n(L, l_2, l_3) T_{l_3 l_4}^{l_1 l_2}(L), \end{aligned} \quad (73)$$

This term will contribute at a similar order to the third term with  $\langle \bar{\beta}_n^{\kappa^2} \rangle_3 \sim F_{\text{NL}}^2$  if there is a significant trispectrum, that is, if  $g_{\text{NL}} \tau_{\text{NL}} \gg 1$  (see ref. [44]). Assuming that correlators (unconnected parts) beyond fourth-order are negligible, then the primary contributions to  $\bar{F}_{\text{NL}}^2$  become

$$\bar{F}_{\text{NL}}^2 = 6n_{\max} + \sum_n^{n_{\max}} [F_{\text{NL}}^2 \bar{\alpha}_n^{\kappa^2} + \langle \bar{\beta}_n^{\kappa^2} \rangle_6 + \dots]. \quad (74)$$

- 
- [1] J. R. Fergusson, M. Liguori, and E. P. S. Shellard (2009), 0912.5516.
  - [2] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolte, L. Page, et al., ArXiv e-prints (2010), 1001.4538.
  - [3] M. Liguori, E. Sefusatti, J. R. Fergusson, and E. P. S. Shellard, ArXiv e-prints (2010), 1001.4707.
  - [4] J. R. Fergusson and E. P. S. Shellard, Phys. Rev. **D76**, 083523 (2007), astro-ph/0612713.
  - [5] J. R. Fergusson and E. P. S. Shellard, Phys. Rev. **D80**, 043510 (2009), 0812.3413.
  - [6] D. Munshi and A. Heavens (2009), 0904.4478.
  - [7] M. Bucher, B. Van Tent, and C. S. Carvalho (2009), 0911.1642.
  - [8] X.-c. Luo, Astrophys. J. **427**, L 71 (1994), astro-ph/9312004.
  - [9] E. Komatsu and D. N. Spergel, Phys. Rev. **D63**, 063002 (2001), astro-ph/0005036.
  - [10] E. Komatsu, D. N. Spergel, and B. D. Wandelt, Astrophys. J. **634**, 14 (2005), astro-ph/0305189.
  - [11] P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark, and M. Zaldarriaga, JCAP **0605**, 004 (2006), astro-ph/0509029.
  - [12] K. M. Smith, L. Senatore, and M. Zaldarriaga, JCAP **0909**, 006 (2009), 0901.2572.
  - [13] E. Komatsu et al. (WMAP) (2008), 0803.0547.
  - [14] A. P. S. Yadav and B. D. Wandelt, Phys. Rev. Lett. **100**, 181301 (2008), 0712.1148.
  - [15] J. R. Fergusson, M. Liguori, and E. P. S. Shellard (2010).
  - [16] M. Liguori, S. Matarrese, and L. Moscardini, Astrophys. J. **597**, 57 (2003), astro-ph/0306248.
  - [17] E. Komatsu, A. Kogut, M. R. Nolte, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, M. Limon, S. S. Meyer, L. Page, et al., Astrophys. J. **148**, 119 (2003), arXiv:astro-ph/0302223.
  - [18] D. N. Spergel et al. (WMAP), Astrophys. J. Suppl. **170**, 377 (2007), astro-ph/0603449.
  - [19] X. Chen and Y. Wang (2009), 0909.0496.
  - [20] J. M. Maldacena, JHEP **05**, 013 (2003), astro-ph/0210603.
  - [21] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Nucl. Phys. **B667**, 119 (2003), astro-ph/0209156.
  - [22] G. I. Rigopoulos, E. P. S. Shellard, and B. J. W. van Tent, Phys. Rev. **D73**, 083522 (2006), astro-ph/0506704.
  - [23] D. Seery and J. E. Lidsey, JCAP **0506**, 003 (2005), astro-ph/0503692.

- [24] F. Vernizzi and D. Wands, JCAP **5**, 19 (2006), arXiv:astro-ph/0603799.
- [25] A. Linde and V. Mukhanov, JCAP **0604**, 009 (2006), astro-ph/0511736.
- [26] D. H. Lyth, C. Ungarelli, and D. Wands, Phys. Rev. **D67**, 023503 (2003), astro-ph/0208055.
- [27] N. Bartolo, S. Matarrese, and A. Riotto, Phys. Rev. **D69**, 043503 (2004), hep-ph/0309033.
- [28] K. Enqvist, A. Jokinen, A. Mazumdar, T. Multamaki, and A. Vaihkonen, Phys. Rev. Lett. **94**, 161301 (2005), astro-ph/0411394.
- [29] N. Barnaby and J. M. Cline, JCAP **0806**, 030 (2008), 0802.3218.
- [30] J.-L. Lehners and P. J. Steinhardt, Phys. Rev. **D77**, 063533 (2008), 0712.3779.
- [31] X. Chen (2010), 1002.1416.
- [32] M. Alishahiha, E. Silverstein, and D. Tong, Phys. Rev. **D70**, 123505 (2004), hep-th/0404084.
- [33] X. Chen, R. Easther, and E. A. Lim, JCAP **0706**, 023 (2007), astro-ph/0611645.
- [34] P. Creminelli, JCAP **0310**, 003 (2003), astro-ph/0306122.
- [35] N. Arkani-Hamed, P. Creminelli, S. Mukohyama, and M. Zaldarriaga, JCAP **0404**, 001 (2004), hep-th/0312100.
- [36] P. D. Meerburg, J. P. van der Schaar, and P. S. Corasaniti, JCAP **0905**, 018 (2009), 0901.4044.
- [37] R. Holman and A. J. Tolley, JCAP **0805**, 001 (2008), 0710.1302.
- [38] I. G. Moss and C. Xiong, JCAP **0704**, 007 (2007), astro-ph/0701302.
- [39] A. Berera, I. G. Moss, and R. O. Ramos, Rept. Prog. Phys. **72**, 026901 (2009), 0808.1855.
- [40] I. G. Moss and C. M. Graham, JCAP **0711**, 004 (2007), 0707.1647.
- [41] X. Chen, R. Easther, and E. A. Lim, JCAP **0804**, 010 (2008), 0801.3295.
- [42] R. Bean, X. Chen, G. Hailu, S. H. H. Tye, and J. Xu, JCAP **0803**, 026 (2008), 0802.0491.
- [43] J. R. Fergusson and E. P. S. Shellard, Phys. Rev. **D76**, 083523 (2007), astro-ph/0612713.
- [44] D. M. Regan, E. P. S. Shellard, and J. R. Fergusson (2010), 1004.2915.